

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття освітнього ступеня

бакалавр

(назва освітнього ступеня)

на тему: *Комп'ютерна система автоматичного розпізнавання та класифікації звукових сигналів*

Виконав: студент IV курсу, групи СІс-43
спеціальності 123 «Комп'ютерна інженерія»

(шифр і назва спеціальності)

(підпис)

Сахарчук В.В.

(прізвище та ініціали)

Керівник

(підпис)

Яцишин В.В.

(прізвище та ініціали)

Нормоконтроль

(підпис)

Луцик Н.С.

(прізвище та ініціали)

Завідувач кафедри

(підпис)

Осухівська Г.М.

(прізвище та ініціали)

Рецензент

(підпис)

Бойко І.В.

(прізвище та ініціали)

Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя

Факультет комп'ютерно-інформаційних систем і програмної інженерії
(повна назва факультету)

Кафедра комп'ютерних систем та мереж
(повна назва кафедри)

ЗАТВЕРДЖУЮ
Завідувач кафедри
Осухівська Г.М.
(підпис) (прізвище та ініціали)
« » 2022 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

на здобуття освітнього ступеня бакалавр
(назва освітнього ступеня)

за спеціальністю 123 «Комп'ютерна інженерія»
(шифр і назва спеціальності)

студенту Сахарчуку Вадиму Вікторовичу
(прізвище, ім'я, по батькові)

1. Тема роботи Комп'ютерна система автоматичного розпізнавання та класифікації звукових сигналів

Керівник роботи Яцишин Василь Володимирович, к.т.н., доцент
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

Затверджені наказом ректора від «23» березня 2022 року № 4.7-180

2. Термін подання студентом завершеної роботи 24.06.2022 р.

3. Вихідні дані до роботи Методи представлення звукового сигналу, параметри звукового сигналу, методи розпізнавання звукових хвиль

4. Зміст роботи (перелік питань, які потрібно розробити)

Вступ. 1. Аналіз вимог та особливостей побудови систем розпізнавання і класифікації звукових сигналів. 2. Проектування апаратної складової системи розпізнавання та класифікації звукових сигналів. 3. Реалізація програмної моделі розпізнавання і класифікації звукових сигналів. 4. Безпека життєдіяльності, основи охорони праці. Висновки

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень, слайдів)

1. Загальна характеристика звукових сигналів

2. Схема електрична принципова Arduino Nano 33 BLE Sense

3. Алгоритм реалізації програмного забезпечення класифікації звукових сигналів.

4. Спектрограма звукових сигналів автомобілів екстрених служб та звичайних авто.

5. Структура моделі та результати класифікації звукових сигналів

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
<i>Безпека життєдіяльності, основи охорони праці</i>	<i>Лазарюк В.В., к.т.н., доц. каф. МТ</i>		

7. Дата видачі завдання _____

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1	<i>Розробка та аналіз технічного завдання</i>	<i>23.03-30.03.2022</i>	
2	<i>Аналіз принципів організації комп'ютерних систем розпізнавання і класифікації звукових сигналів</i>	<i>30.03-06.04.2022</i>	
3	<i>Обґрунтування вибору апаратного забезпечення комп'ютерної системи</i>	<i>07.04-15.04.2022</i>	
4	<i>Проектування архітектури системи розпізнавання і класифікації звукових сигналів</i>	<i>15.04-29.04.2022</i>	
5	<i>Аналіз параметрів звукових сигналів у цифровому форматі</i>	<i>30.04-12.05.2022</i>	
6	<i>Реалізація програмної моделі інтелектуального модуля</i>	<i>13.05-18.05.2022</i>	
7	<i>Розробка інструкцій з налаштування параметрів комп'ютерної системи</i>	<i>18.05-25.05.2022</i>	
8	<i>Безпека життєдіяльності, основи охорони праці</i>	<i>25.05-10.06.2022</i>	
9	<i>Оформлення кваліфікаційної роботи</i>	<i>10.06-12.06.2022</i>	
10	<i>Попередній захист кваліфікаційної роботи</i>	<i>13.06-18.06.2022</i>	
11	<i>Захист кваліфікаційної роботи</i>	<i>22.06-24.06.2022</i>	

Студент

_____ (підпис)

Сахарчук Вадим Вікторович

_____ (прізвище та ініціали)

Керівник роботи

_____ (підпис)

Яцишин Василь Володимирович

_____ (прізвище та ініціали)

АНОТАЦІЯ

Комп'ютерна система автоматичного розпізнавання та класифікації звукових сигналів // Кваліфікаційна робота на здобуття освітнього ступеня бакалавр // Сахарчук Вадим Вікторович// ТНТУ, спеціальність 123 «Комп'ютерна інженерія»// Тернопіль, 2022 // с.– 86, рис. – 32 , табл. – 3, аркушів А1 – 6, бібліогр. – 19.

Ключові слова: розпізнавання, звук, сигнал, класифікація, система.

При виконанні кваліфікаційної роботи досягнуто поставленої мети щодо проектування комп'ютерної системи автоматичного розпізнавання і класифікації звукових сигналів на основі методів і засобів спектрального аналізу дискретних звукових сигналів та параметрів його амплітуди і частоти.

Комп'ютерна система на апаратному рівні використовує плату Arduino Nano 33 BLE Sense, що містить в собі ряд інтегрованих сенсорів, серед яких мікрофон підвищеної чутливості з шумопоглинанням. Це дає змогу одержувати звуковий сигнал в реальному часі і за допомогою реалізованої моделі інтелектуального розпізнавання і класифікації аудіо сигналів розпізнавати звуки сирен автомобілів екстрених служб.

Інтелектуальну систему розпізнавання і класифікації звукових сигналів реалізовано за допомогою мови програмування Python та відкритих бібліотек. Після цього дану модель завантажено в оперативну пам'ять Arduino Nano. Точність розпізнавання звукових сигналів становить вище 90%, а у деяких випадках досягає 97%.

ABSTRACT

Computer system for sound signals automatic recognition and classification // Bachelor's thesis // Sakharchuk Vadym Viktorovych // TNTU, speciality 123 «Computer engineering»// Ternopil, 2022 // p.– 86 , fig. – 32 , tab. – 3, posters A1 – 6, ref. – 19.

Keywords: recognition, sound, signal, classification, system.

The qualification work achieved the goal of designing a computer system for automatic recognition and classification of sound signals based on methods and means of spectral analysis of discrete sound signals and parameters of its amplitude and frequency.

The computer-level computer system uses an Arduino Nano 33 BLE Sense board, which includes a number of integrated sensors, including a high-sensitivity microphone with noise absorption. This makes it possible to receive a sound signal in real time and with the help of the implemented model of intelligent recognition and classification of audio signals to recognize the sounds of sirens of emergency vehicles.

The intelligent sound recognition and classification system is implemented using the Python programming language and open libraries. This model is then loaded into Arduino Nano RAM. The accuracy of sound signal recognition is above 90%, and in some cases reaches 97%.

ЗМІСТ

ПЕРЕЛІК ОСНОВНИХ УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ І СКОРОЧЕНЬ	8
ВСТУП	9
РОЗДІЛ 1 АНАЛІЗ ВИМОГ ТА ОСОБЛИВОСТЕЙ ПОБУДОВИ СИСТЕМ РОЗПІЗНАВАННЯ І КЛАСИФІКАЦІЇ ЗВУКОВИХ СИГНАЛІВ	10
1.1 Аналіз технічного завдання на проектування комп'ютерної системи прогнозування ризиків появи серцевих захворювань.....	10
1.2 Аналіз особливостей, параметрів та застосування аналізу звукових сигналів	13
РОЗДІЛ 2 ПРОЕКТУВАННЯ АПАРАТНОЇ СКЛАДОВОЇ СИСТЕМИ РОЗПІЗНАВАННЯ ТА КЛАСИФІКАЦІЇ ЗВУКОВИХ СИГНАЛІВ	22
2.1 Застосування мікроконтролерів та систем машинного навчання	22
2.2 Особливості і технічні характеристики Arduino Nano 33 BLE Sense	25
2.3 Характеристики сенсора-мікрофона MP34DT05-A.....	31
2.4 Аналіз природи звукових сигналів	34
РОЗДІЛ 3 РЕАЛІЗАЦІЯ ПРОГРАМНОЇ МОДЕЛІ РОЗПІЗНАВАННЯ І КЛАСИФІКАЦІЇ ЗВУКОВИХ СИГНАЛІВ	42
3.1 Побудова процедури реалізації програмного забезпечення аналізу звукових сигналів і препроцесинг даних	42
3.2 Побудова моделі машинного навчання на основі ознак часової області ...	50
3.3 Модель класифікації з врахуванням ознак спектрограми.....	58
РОЗДІЛ 4 БЕЗПЕКА ЖИТТЄДІЯЛЬНОСТІ, ОСНОВИ ОХОРОНИ ПРАЦІ	65

					КС КРБ 123.232.00.00 ПЗ			
Змн.	Арк.	№ докум.	Підпис	Дата	Комп'ютерна система автоматичного розпізнавання та класифікації звукових сигналів	Літ.	Арк.	Аркуші
Розроб.		Сахарчук В.В.					6	
Перевір.		Яцишин В.В.				ТНТУ, каф. КС, гр. СІс-43		
Реценз.								
Н. Контр.		Луцик Н.С.						
Затверд.		Осухівська Г.М.						

4.1	Вимоги до режимів праці і відпочинку при роботі з ВДТ	65
4.2	Вплив електромагнітних полів (ЕМП) на людину та заходи щодо зменшення їх впливу на обслуговуючий персонал	67
	ВИСНОВКИ	72
	СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	73
Додаток А. Технічне завдання		

					КС КРБ 123.232.00.00 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		7

ПЕРЕЛІК ОСНОВНИХ УМОВНИХ ПОЗНАЧЕНЬ,
СИМВОЛІВ І СКОРОЧЕНЬ

БД	База даних
КС	Комп'ютерна система
ПЗ	Програмне забезпечення
ML	Machine Learning
IEEE	Institution of Electrical and Electronics Engineers
ISO	International Organization for Standartization

					<i>КС КРБ 123.232.00.00 ПЗ</i>	Арк.
						8
Змн.	Арк.	№ докум.	Підпис	Дата		

ВСТУП

Потужний стрибок інформаційних технологій ХХІ стимулює розвиток все нових сфер їхнього застосування, що покликаний автоматизувати рутинні операції та підвищити зручність і ефективність виробничих процесів та надання послуг.

На сьогодні можна спостерігати масштабну імплементацію методів і засобів штучного інтелекту, а також IoT-технологій у повсякденне життя людей. Зараз вже нікого не здивуєш голосовим управлінням побутовими пристроями, автомобілями та іншими зручними гаджетами. Це все об'єднує нову технологію під назвою «розумний будинок», або ще ширше «розумні технології».

Важливих здобутків досягнуто також у сферах розпізнавання образів, опрацювання природної мови та ряду інших, що дозволяє покращити і підвищити якість життя населення. Однак визначальним фактором у життєдіяльності людини відіграє якість і вчасність надання послуг екстреними службами у разі виникнення аварійних ситуацій, зокрема, швидкими допомогоюми, пожежниками, поліцією та ін.

Враховуючи шалені темпи розвитку дорожнього трафіку, пов'язаного із збільшенням автотранспорту та повільні кроки до розширення інфраструктури у великих містах виникають проблеми із заторами автомобілів.

Наявність заторів у містах не дозволяє ефективно працювати екстреним службам, оскільки накопичення авто на світлофорах та недосконала інфраструктура не дають можливості швидко прибувати до місця виникнення надзвичайної ситуації. Тому актуальною задачею є розробка комп'ютерної системи автоматичного розпізнавання і класифікації звукових сигналів, які генеруються сиренами автомобілів спеціального призначення, для формування зелених коридорів при проїзді перехресть зі світлофорним регулюванням.

					КС КРБ 123.232.00.00 ПЗ	Арк.
						9
Змн.	Арк.	№ докум.	Підпис	Дата		

РОЗДІЛ 1 АНАЛІЗ ВИМОГ ТА ОСОБЛИВОСТЕЙ ПОБУДОВИ СИСТЕМ РОЗПІЗНАВАННЯ І КЛАСИФІКАЦІЇ ЗВУКОВИХ СИГНАЛІВ

1.1 Аналіз технічного завдання на проектування комп'ютерної системи прогнозування ризиків появи серцевих захворювань

Комп'ютерна система розпізнавання та класифікації звукових сигналів призначена для автоматизації процесу визначення і встановлення приналежності згенерованих джерелом звуку сигналів на основі яких приймаються певні рішення.

Проектована система повинна забезпечувати можливість збору звукових сигналів і перетворення їх у цифровий формат. Окрім цього, одержані сигнали з відповідними характеристиками на основі моделей та алгоритмів машинного навчання повинні бути автоматично прокласифіковані за відповідними класами з відповідними мітками.

Комп'ютерні системи розпізнавання і класифікації звукових сигналів мають широкий спектр застосування, починаючи від автономних незалежних систем, закінчуючи підсистемами більш комплексних ІТ-рішень.

Автономно такі системи можуть бути використані при формуванні альбомів музики за жанрами, виконавцями чи іншими категоріями. В якості підсистем більш складних рішень, комп'ютерна система розпізнавання і класифікації звукових сигналів може бути використана при організації безпеки та голосового управління «розумного будинку».

Окрім цього, зважаючи на кількість автомобілів і їх постійне зростання при незмінній інфраструктурі, актуальним є створення систем аналізу звукових сигналів автомобілів. Це дозволить екстреним службам більш швидко та

					КС КРБ 123.232.00.00 ПЗ			
<i>Змн.</i>	<i>Арк.</i>	<i>№ докум.</i>	<i>Підпис</i>	<i>Дата</i>				
<i>Розроб.</i>		Сахарчук В.В.			<i>Аналіз вимог та особливостей побудови систем розпізнавання і класифікації звукових сигналів</i>	<i>Лім.</i>	<i>Арк.</i>	<i>Аркушів</i>
<i>Перевір.</i>		Яцишин В.В.					10	
<i>Реценз.</i>						<i>ТНТУ, каф. КС, гр. СІс-43</i>		
<i>Н. Контр.</i>		Луцик Н.С.						
<i>Затверд.</i>		Осухівська Г.М.						

ефективно добиратись до місць виникнення небезпечних ситуацій, шляхом формування зелених коридорів.

Комп'ютерна система повинна дозволяти накопичувати дані про звукові сигнали автомобілів та автоматично встановлювати мітки приналежності звукового сигналу. На апаратному рівні система повинна бути швидкореагуючою на події, переносимою і здатною співіснувати з іншими сумісними системами.

Мета проектування та реалізації комп'ютерної системи розпізнавання та класифікації звукових сигналів полягає у створенні ефективного рішення як на апаратному, так і на програмному рівні, яка б дозволяла проводити бінарну класифікацію і прогнозування звукових сигналів для екстрених служб і формування пріоритету їх проїзду на перехрестях із світлофорами.

Передбачається, що досягнення мети роботи можливе при розв'язанні наступних задач:

- аналіз природи і параметрів звукових сигналів;
- аналіз методів і засобів перетворення аналогових звукових сигналів у цифровий формат;
- обґрунтування апаратного забезпечення при проектуванні комп'ютерної системи;
- аналіз програмних засобів для роботи із звуковими сигналами та моделей машинного навчання для їх автоматичного розпізнавання та класифікації;
- реалізація програмних моделей класифікації аудіосигналів з визначеною точністю і достовірністю;
- аналіз результатів проектування і реалізації комп'ютерної системи.

До основних задач і функцій системи розпізнавання і класифікації звукових сигналів належить здатність до накопичення і зберігання аудіо даних, автоматичної класифікації сигналів за визначеними класами і нанесення відповідних міток на прогнозовані сигнали.

					КС КРБ 123.232.00.00 ПЗ	Арк.
						11
Змн.	Арк.	№ докум.	Підпис	Дата		

Реалізацію комп'ютерної системи необхідно виконати із застосуванням мікроконтролерів з підтримкою та опрацювання алгоритмів машинного навчання. Продуктивність системи повинна бути високою для забезпечення зелених коридорів спецмашинам у заторах та на перехрестях зі світлофорами.

Система повинна володіти характеристиками щодо мінімальності габаритних розмірів і високої чутливості мікрофона з можливістю подавлення зовнішніх шумів.

Найбільш критичними задачами, які повинна розв'язувати комп'ютерна система є автоматична класифікація звукових сигналів у визначених часових рамках. Окрім цього, точність розпізнавання сирен автомобілів екстрених служб повинна бути на рівні не менше, ніж 90 %.

Додатковими задачами, які потрібно вирішити поряд з основною є наступні:

- апаратна та програмна узгодженість компонентів комп'ютерної системи;
- наявність можливості функціонування моделі класифікації звукових сигналів на мобільних і портативних пристроях, в основі якого лежить мікроконтролер;
- застосування звукових даних, які інтерпретують сирени спецавтомобілів, з відкритих джерел;
- використання швидких та робастних алгоритмів при бінарній класифікації звуків сирен;
- забезпечення гнучкості налаштування параметрів при навчанні моделі розпізнавання звукових сигналів.

Найбільш раціональним є використання апаратного забезпечення з вбудованими сенсорами звуку.

Обґрунтування вибору мікрофону зумовлено факторами його розміру, типом з'єднання з мікроконтролером та підвищеною чутливістю з шумопоглинальним фільтром.

					КС КРБ 123.232.00.00 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		12

Вимогами до мікроконтролера є достатність системних параметрів, зокрема тактової частоти процесора, його розрядності та оперативної пам'яті для швидкої видачі результатів щодо одержаного звукового сигналу в режимі реального часу.

Програмна складова комп'ютерної системи повинна забезпечувати сумісність з особливостями апаратного забезпечення та зручність використання та налаштування моделей розпізнавання і класифікації звукових сигналів.

Основні вимоги до комп'ютерної системи:

- можливість одержання аналогового звукового сигналу і перетворення його у цифровий;
- наявність аудіо файлів із записами звукових сигналів сирен автомобілів екстрених служб;
- можливість навчання та донавчання моделі розпізнавання і класифікації звукових сигналів;
- здатність налаштування чутливості приймача звукових сигналів;
- здатність до інтеграції з системами керування світлофорами;
- забезпечення надійності та захищеності апаратного і програмного забезпечення формування міток звукових сигналів.

Вимогами до способів і засобів зв'язку є здатність здійснювати обмін даними за допомогою енергозберігаючих протоколів Bluetooth у випадку взаємодії із зовнішнім ПК або суміжними системами. Сама комп'ютерна система на рівні апаратного забезпечення повинна бути монолітною.

1.2 Аналіз особливостей, параметрів та застосування аналізу звукових сигналів

Мовлення визначається як вираження думок і почуттів за допомогою артикуляції звуків. Воно є найбільш природним, інтуїтивним і найулюбленішим засобом спілкування людей.

					КС КРБ 123.232.00.00 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		13

Перцептивна мінливість мовлення існує у вигляді різних мов, діалектів, наголосів, а словниковий запас мовлення з кожним днем збільшується.

Більш складна мінливість на рівні мовного сигналу існує у вигляді різної амплітуди, тривалості, висоти, тембру та джерела звуку.

Текст як засіб комунікації розвивався для зберігання та передачі інформації на великі відстані. Це письмове відображення будь-якого мовленнєвого спілкування. Це більш проста форма спілкування і позбавлена вищезгаданих складних варіацій, що існують у мовленні.

Складні мінливості мовлення ускладнюють аналіз, але надають додаткову інформацію, використовуючи зміну тону та амплітуди.

Аналіз мовлення та тексту має широке застосування в сучасному світі. Вони мають різні представлення, і в їх аналізі зустрічається багато відмінностей і проблем.

Розглянемо звуковий сигнал, взятий з одного висловлювання бази даних CMU US RMS ARCTIC. Тут кожне з висловлювань записується як 16-бітовий сигнал з частотою дискретизації 16 кГц, що означає, що на кожну секунду сигналу припадає 16000 вибірок, і кожна вибірка має роздільну здатність 16 біт на вибірку.

Частота дискретизації аудіосигналу визначає роздільну здатність аудіосемплів, чим вище частота дискретизації, тим вище роздільна здатність сигналу.

Звуковий сигнал зчитується з файлу «arctic_a0005.wav» бази даних розмовного мовлення, тривалість якого становить близько 1,4 секунди, що еквівалентно послідовності з 22640 вибірок, кожна вибірка має 16-бітове число.

Нижче наведене звукове представлення (рисунок 1.1), що є відображенням звукового сигналу з «arctic_a0005.wav» та відповідає тексту: «чи ми коли-небудь забудемо це»:

					КС КРБ 123.232.00.00 ПЗ	Арк.
						14
Змн.	Арк.	№ докум.	Підпис	Дата		

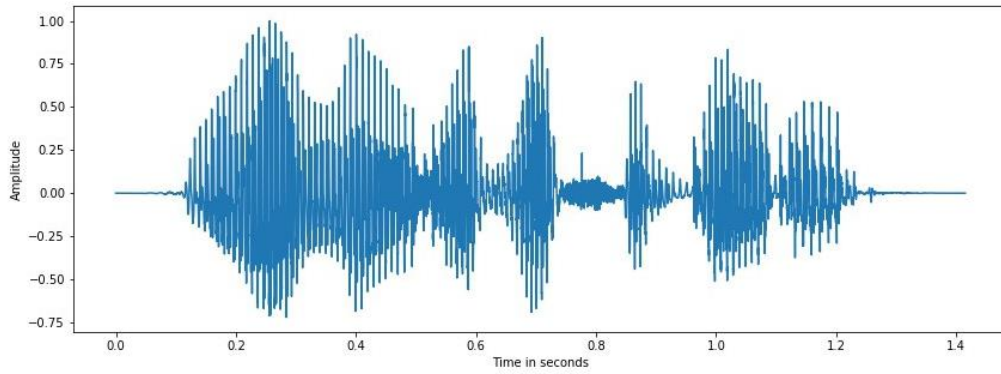


Рисунок 1.1 – Відображення звукового сигналу

З наведеного вище рисунка (сигнал $s[n]$ для висловлювання «чи забудемо ми це коли-небудь») видно, що звуковий сигнал може бути представлений як зміна амплітуди в часі. Амплітуда нормується таким чином, що максимальне значення дорівнює 1.

Мовлення – це в основному послідовність артикуляційних звукових одиниць, таких як «w», «ih», відомих як фонема.

Мовний сигнал можна розділити на послідовність фонем і тих чи інших немовних сегментів. База даних також містить відповідну текстову транскрипцію для мовного сигналу як на рівні речення, так і на рівні фонем для кожного файлу звукової хвилі.

Нижче (рисунок 1.2) наведено представлення частини вищевказаного мовного сигналу, що показує фонем та відповідний їм часовий проміжок.

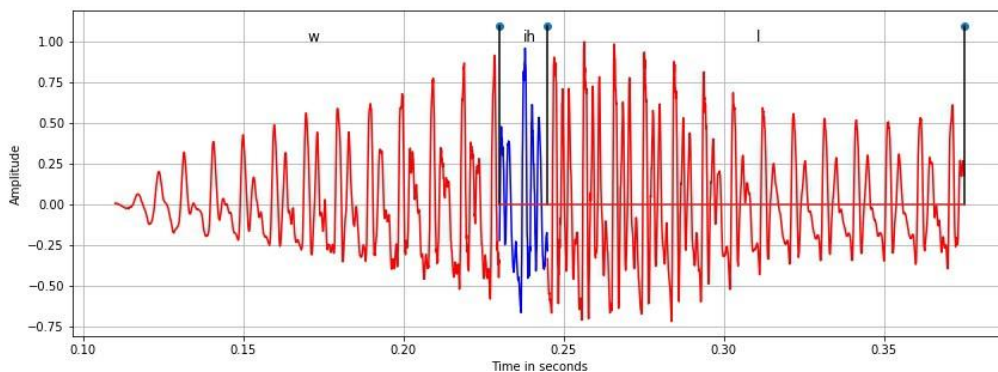


Рисунок 1.2 – Фонем у звуковому сигналі

З рисунка 1.2 видно, що фонemi «w», «ih» та «l» мають квазіперіодичну природу і класифікуються як дзвінкі фонemi, оскільки вони утворені періодичною вібрацією голосових зв'язок.

Крім того, «ih» є голосним, тоді як «w» і «l» є напівголосними. Класи дзвінких і глухих – це широка категорія звуків мови, заснована на вібрації голосових зв'язок (рисунок 1.3). Вивчення та класифікація різних фонем називається фонетикою.

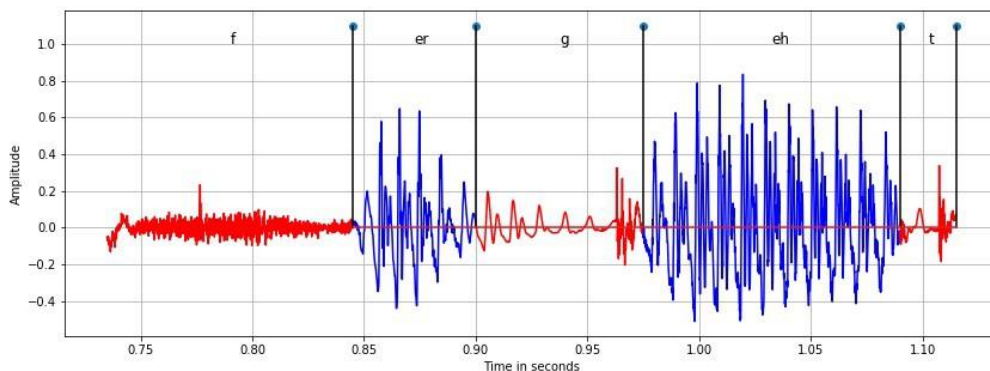


Рисунок 1.3 – Приклади різних видів фонем

На рисунку 1.3 показано глухі фонemi, такі як «f», «g» і «t», і дзвінкі фонemi, такі як «er» і «eh». Фонemi «g» і «t» далі класифікуються як зупинки, що означає мовчання, за яким слідує раптовий імпульс.

Можна помітити, що дзвінкий компонент є квазіперіодичним, тоді як глухий є шумним, оскільки не утворюється періодичною вібрацією голосових зв'язок. Фонemi можна зіставити з письмовою формою мови за допомогою зіставлення фонemi та графем. На рисунку 1.4 наведено порівняння між текстом і відповідними фонемами:

Текст:

«will we ever forget it»



Фонетична послідовність:

"w", "ih", "l", "w", "iy", "eh", "v", "er", "f", "er", "g", "eh", 't', 'ih', 't'

Рисунок 1.4 – Порівняння тексту з фонетичним його відображенням

З наведеного вище зіставлення видно, що слово «воля» співставляється з фонемами «w», «ih», «l».

Далі розглянемо представлення мови в частотній області кадру, також відоме як швидке перетворення Фур'є у дослідницькій спільноті мовлення. Спектрограма – це візуальне уявлення звуку в частотній області. Приклад спектрограми показано на рисунку 1.5.

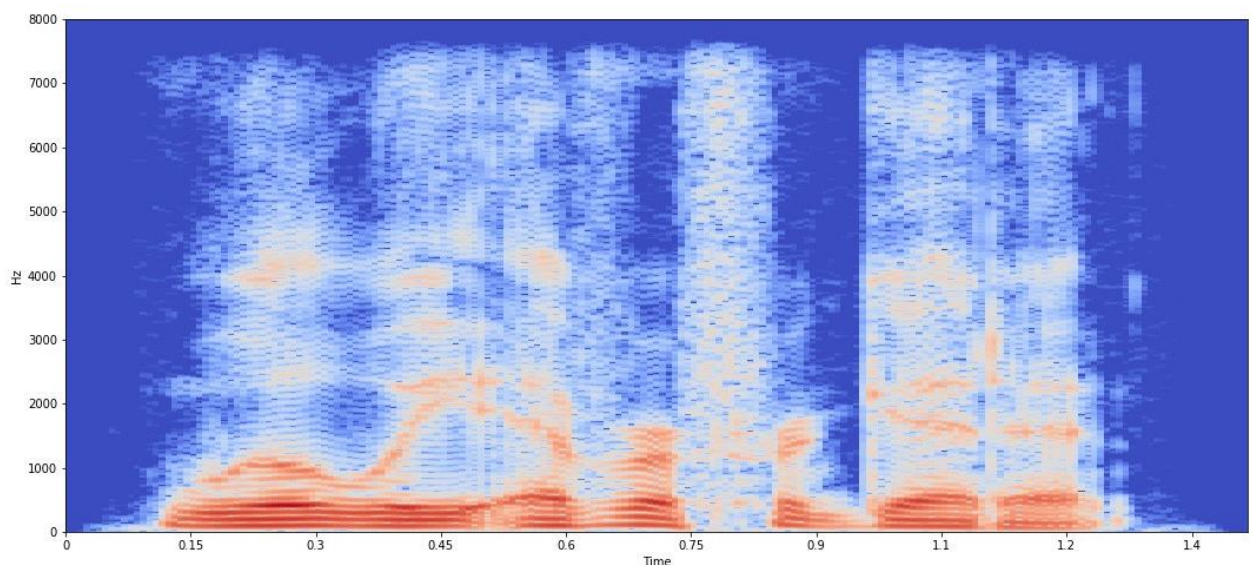


Рисунок 1.5 – Приклад спектрограми

Логарифмічно масштабована спектрограма, наведена вище, є амплітудою STFT у логарифмічній шкалі. Вибирається розмір кадру 30 мс, що еквівалентно

30 мс*16 кГц = 480 вибірок звукового сигналу, тоді як зсув кадру 7,5 мс еквівалентний 120 вибіркам.

Причина поділу звукового сигналу на кадри малої тривалості полягає в тому, що він є нестационарним і його тимчасові характеристики дуже швидко змінюються.

Отже, взявши малий розмір кадру, можна припустити, що звуковий сигнал буде нерухомим і його характеристики не будуть сильно змінюватися в межах кадру.

Крім того, вибирається коротший зсув кадру, щоб відстежувати безперервність сигналу і не пропускати жодних різких змін на краях кадрів.

З наведеного вище графіка видно, що представлення в частотній області кожного кадру допомагає краще аналізувати звуковий сигнал; оскільки можна легко побачити гармоніки як паралельні червоні сегменти в голосових областях і як змінюється амплітуда для кожної частоти та індексу кадру.

Отже, більшість аналізу звукового сигналу виконується в частотній області. Але вилучення тимчасової інформації, як-от різкі зміни в сигналі (початок пакету, як «t»), краще фіксується у часовій області, оскільки поділ сигналу на кадри відкидає миттєві зміни в сигналі. Можна сказати, що для отримання кращої тимчасової роздільної здатності у частотній області, необхідно взяти менший розмір кадру. Але існує компроміс між роздільною здатністю в часовій і частотній областях.

Взяття дуже малого розміру кадру дасть вищу роздільну здатність за часом, але наявними будуть кілька вибірок в одному кадрі, а відповідні компоненти Фур'є матимуть кілька частотних компонентів. Використання більшого розміру кадру дасть меншу тимчасову роздільну здатність, але вищу роздільну здатність за частотою через більшу кількість вибірок. Таким чином, отримати високу роздільну здатність одночасно за часом і частотою неможливо.

Помічено, що вісь ординат на діаграмі логарифмічної спектрограми має частоту до 8 кГц для частоти дискретизації 16 кГц. Це пояснюється тим, що, згідно з теоремою дискретизації Найквіста-Шеннона, максимальна частота, яку

					КС КРБ 123.232.00.00 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		18

можна спостерігати в дискретному сигналі, становить не більше половини частоти дискретизації, яка становить 8 кГц.

У той час як мова має багато змінних параметрів залежно від середовища, мовця, настрою та тону, текст позбавлений усієї цієї мінливості.

Еквівалентний текст представлений у вигляді послідовності алфавітів, символів і пробілів «чи забудемо ми це коли-небудь». Сферами застосування парадигм аналізу звукових сигналів є:

- виявлення голосової активності – визначення сегментів у формі звукового сигналу, де присутня лише мова, ігноруючи немовні та беззвучні сегменти;

- покращення мовлення – покращення якості звукового сигналу шляхом фільтрації та відділення шуму від звукових сегментів;

- розпізнавання мовлення – перетворення звукового сигналу у текст все ще є проблемою в різних умовах, розпізнавання може залежати від словникового запасу або не залежати від нього;

- перетворення тексту у звук – синтезувати природне мовлення з тексту, щоб мова звучала дуже природно з емоціями є складним завданням;

- діаризація та розпізнавання джерела звуку – діаризація представляє собою сегментацію звукового сигналу, що належать різним особам, тоді як розпізнавання джерела звуку визначає, хто говорить у певний час;

- розділення джерел звуку – розділення змішаного звукового сигналу, наприклад мовлення, яке перекривається звуками іншого динаміка або шуму;

- модифікація мовлення – модифікація мови, наприклад зміна її емоцій, тону, перетворення на мовлення іншої особи;

- класифікація емоційного мовлення – визначення емоцій, таких як радісний, сердитий, сумний і тривожний.

Класифікація аудіо – це процес прослуховування та аналізу аудіозаписів. Цей процес, також відомий як класифікація звуку, лежить в основі різноманітних сучасних технологій штучного інтелекту, включаючи віртуальних помічників, автоматичне розпізнавання мовлення та програми, які перетворюють текст в

					КС КРБ 123.232.00.00 ПЗ	Арк.
						19
Змн.	Арк.	№ докум.	Підпис	Дата		

мовлення. Також класифікацію звукових сигналів можна знайти у сферах прогнозування технічного обслуговування, системах безпеки розумного дому, індексації та пошуку мультимедіа.

Проекти класифікації аудіо, подібні до згаданих вище, починаються з анотованих аудіоданих. Машинам потрібні ці дані, щоб навчитися чути і слухати. Використовуючи ці дані, вони розвивають здатність розрізняти звуки для виконання конкретних завдань.

Процес анотації часто включає класифікацію аудіофайлів на основі конкретних потреб проекту за допомогою спеціальних служб класифікації аудіо. Нижче наведено чотири типи класифікації та відповідні варіанти використання для кожного.

1. Класифікація акустичних даних. Цей тип класифікації також відомий як виявлення акустичних подій, який визначає, де був записаний звуковий сигнал. Це означає розмежування між такими середовищами, як ресторани, школи, будинки, офіси, вулиці тощо. Одним із застосуванням класифікації акустичних даних є створення та підтримка звукових бібліотек для аудіо-мультимедіа. Він також відіграє роль у моніторингу екосистеми. Одним із прикладів цього є оцінка чисельності риби в певній частині океану на основі їх акустичних даних.

2. Класифікація звуків навколишнього середовища: як впливає з назви, це класифікація звуків, що зустрічаються в різних середовищах. Наприклад, розпізнавання автомобільних і міських звукових зразків, таких як автомобільні гудки, дорожні роботи, сирени, людські голоси тощо. Це використовується в системах безпеки для виявлення звуків, таких як розбиття скла. Класифікація звукових сигналів також використовується для прогнозування технічного обслуговування пристроїв шляхом виявлення звукових розбіжностей у заводському обладнанні.

3. Класифікація музики. Класифікація музики – це процес визначення категорії до яких належать музика, зокрема, жанр або використовувані музичні інструменти. Ця класифікація відіграє ключову роль в організації аудіобібліотек

					КС КРБ 123.232.00.00 ПЗ	Арк.
						20
Змн.	Арк.	№ докум.	Підпис	Дата		

за жанрами, удосконаленні алгоритмів формування рекомендацій і виявлення тенденцій і переваг слухача за допомогою аналізу даних.

4. Класифікація висловлювань природною мовою: це класифікація записів природною мовою на основі розмовної мови, діалекту, семантики чи інших мовних особливостей. Іншими словами, класифікація людської мови. Цей тип аудіокласифікації найчастіше зустрічається в чат-ботах і віртуальних помічниках, але також поширений у машинному перекладі та програмах синтезу мовлення.

Для проектів, які передбачають класифікацію аудіо, якість набору даних може і буде вирішувати якість результатів проекту. Таким чином, щоб забезпечити точний рівень класифікації аудіо, знадобиться хороший обсяг високоякісних, точно анотованих даних.

TELUS International пропонує набір голосових і звукових даних для машинного навчання, включаючи збір аудіо, транскрипцію та класифікацію. Маючи понад мільйон кваліфікованих працівників, можна збирати різноманітні дані з найрізноманітніших географічних місць і мовних середовищ.

У результаті аналізу звукових сигналів і сфер їх застосування, у кваліфікаційній роботі необхідно реалізувати комп'ютерну систему розпізнавання і класифікації звукових сигналів у сфері виявлення авто екстрених служб.

					КС КРБ 123.232.00.00 ПЗ	Арк.
						21
Змн.	Арк.	№ докум.	Підпис	Дата		

РОЗДІЛ 2 ПРОЕКТУВАННЯ АПАРАТНОЇ СКЛАДОВОЇ СИСТЕМИ РОЗПІЗНАВАННЯ ТА КЛАСИФІКАЦІЇ ЗВУКОВИХ СИГНАЛІВ

2.1 Застосування мікроконтролерів та систем машинного навчання

Arduino ставить перед собою мету зробити машинне навчання досить простим для будь-кого користувача. Сьогодні команда розробників Arduino разом з командою TensorFlow Lite інтегрували TensorFlow Lite Micro в Arduino Nano 33 BLE Sense.

Для реалізації комп'ютерної системи розпізнавання та класифікації звукових сигналів необхідні наступні компоненти:

- плата Arduino Nano 33 BLE Sense;
- кабель Micro USB для підключення плати Arduino до настільного пристрою;
- веб-редактор Arduino або Arduino IDE для програмування плати.

До складу Arduino Nano 33 BLE Sense входять різноманітні вбудовані датчики, що дозволяють будувати хороші додатки з використанням TinyML:

- розпізнавання та аналіз звукових сигналів – цифровий мікрофон;
- аналіз руху об'єктів (object detection) – 9-осьовий IMU (акселерометр, гіроскоп, магнітометр);
- збір та аналіз даних щодо показників навколишнього середовища – температура, вологість і тиск;
- аналіз зображень – яскравість, колір і близькість об'єкта.

На відміну від класичного Arduino Uno, плата поєднує в собі мікроконтролер з вбудованими сенсорами, що означає, що її можна застосовувати у багатьох випадках без додаткового обладнання або зовнішніх пристроїв.

					КС КРБ 123.232.00.00 ПЗ			
<i>Змн.</i>	<i>Арк.</i>	<i>№ докум.</i>	<i>Підпис</i>	<i>Дата</i>				
<i>Розроб.</i>		Сахарчук В.В.			<i>Проектування апаратної складової системи розпізнавання та класифікації звукових сигналів</i>	<i>Лім.</i>	<i>Арк.</i>	<i>Аркушів</i>
<i>Перевір.</i>		Яцишин В.В.					22	
<i>Реценз.</i>						<i>ТНТУ, каф. КС, гр. СІс-43</i>		
<i>Н. Контр.</i>		Луцик Н.С.						
<i>Затверд.</i>		Осухівська Г.М.						

Плата Arduino Nano 33 BLE Sense також досить невелика за габаритами і її можна використовувати як кінцевий мобільний пристрій.

Як випливає з назви, Arduino Nano 33 BLE Sense має підключення Bluetooth® Low Energy, щоб можна було надсилати дані (або результати роботи) на ноутбук, мобільний додаток або інші плати та периферійні пристрої Bluetooth® Low Energy.

Підключення плати BLE Sense через USB – це простий спосіб одержання даних шляхом використання кількох датчиків до одноплатних комп'ютерах без необхідності додаткового підключення обладнання. Це є хорошим доповненням до Raspberry Pi, наприклад, мікроконтролери та TinyML мікроконтролери. Тому такі мікроконтролери використовуються на платах Arduino, є недорогими однокристальними автономними комп'ютерними системами. До них можна віднести «невидимі комп'ютери», вбудовані в мільярди повсякденних гаджетів, таких як мобільні пристрої, дрони, 3D-принтери, іграшки, рисоварки, розумні розетки, електронні скутери, пральні машини.

Тенденція до підключення цих пристроїв є частиною того, що називають Інтернетом речей. Arduino – це платформа та спільнота з відкритим вихідним кодом, зосереджена на тому, щоб зробити розробку програм мікроконтролерів доступною для всіх.

Плата Arduino Nano 33 BLE Sense містить мікроконтролер Arm Cortex-M4, що працює на частоті 64 МГц з 1 МБ флеш-пам'яті та 256 КБ оперативної пам'яті. Це крихітно в порівнянні з хмарою, ПК або мобільним, але розумно за стандартами мікроконтролера. На рисунку 2.1 показано плату Arduino Nano 33 BLE Sense.

					КС КРБ 123.232.00.00 ПЗ	Арк.
						23
Змн.	Арк.	№ докум.	Підпис	Дата		

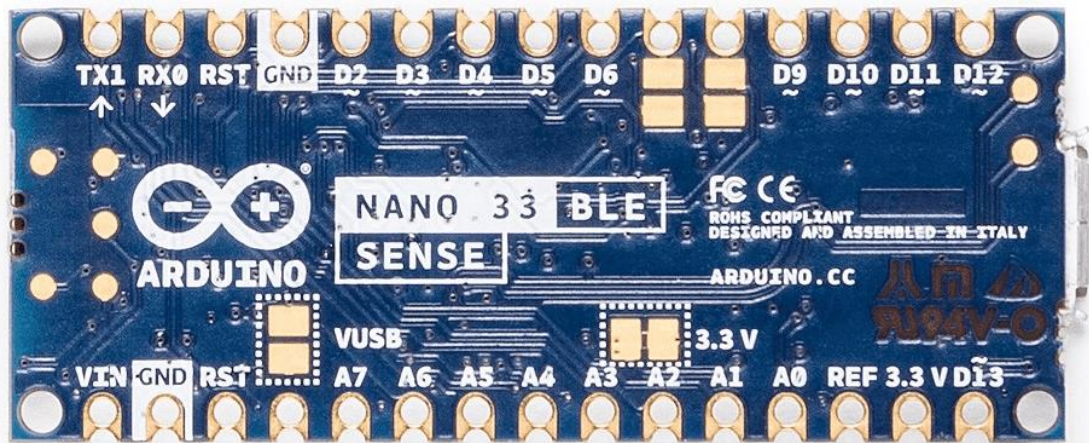


Рисунок 2.1 – Плата Arduino Nano 33 BLE Sense

Плата Arduino Nano 33 BLE Sense за розмірами менша за звичайну гумку. Існують практичні причини, чому варто використовувати ML на мікроконтролерах, зокрема:

- функціональність – «розумний пристрій» забезпечує високу швидкість виконання операцій на локальному рівні (незалежно від Інтернету).№
- вартість – забезпечується простотою використання та не високою вартістю обладнання;
- конфіденційність – забезпечує захист даних і не дозволяє загального доступу ззовні до даних;
- ефективність – проявляється у меншому форм-факторі пристрою, економії енергії та довшому терміну служби акумулятора.

Окрім цього, застосування таких мікроконтролерів є дуже важливим: машинне навчання може зробити мікроконтролери доступними для розробників, які не мають досвіду у розробці вбудованих систем.

Що стосується машинного навчання, існують методи, які можна використовувати, щоб реалізувати моделі нейронної мережі в пристрої з обмеженою пам'яттю, такі як мікроконтролери.

Одним з ключових кроків є квантування ваг від числа з плаваючою комою до 8-бітових цілих чисел. Це також робить свій внесок для пришвидшення обчислень та більшої застосованості з пристроями із нижчою тактовою частотою.

TinyML – це сфера, яка розвивається, і є над чим працювати, але що цікаво, це величезний невивчений простір додатків. Мільярди мікроконтролерів у поєднанні з усіма видами сенсорів у найрізноманітніших місцях, що може привести до серйозних творчих і цінних програм TinyML у майбутньому.

Приклади рішень з використанням TensorFlow Lite для мікроконтролерів тепер упаковані та доступні через диспетчер бібліотек Arduino, що дозволяє включати та запускати їх на Arduino за кілька кліків. Прикладами використання машинного навчання на Arduino є додатки:

- `micro_speech` – розпізнавання мовлення за допомогою вбудованого мікрофона;
- `magic_wand` – розпізнавання жестів за допомогою вбудованого IMU;
- `person_detection` – виявлення людини за допомогою зовнішньої камери ArduCam;

Моделі в цих прикладах були попередньо навчені. Тому при проектуванні комп'ютерної системи розпізнавання і класифікації звукових сигналів можна скористатися попередньо навченими моделями нейронних мереж.

Однак перед тим, як перейти до побудови і налаштування параметрів моделей, необхідно проаналізувати технічну сторону різновиду плати Arduino.

2.2 Особливості і технічні характеристики Arduino Nano 33 BLE Sense

Nano 33 BLE Sense – це плата Arduino з підтримкою моделей штучного інтелекту, яка живиться джерелом постійної напруги 3,3 В і доступна в найменшому форм-факторі 45*18 мм.

Arduino Nano 33 BLE Sense - це абсолютно нова плата у добре відомому форм-факторі, що поставляється з серією вбудованих сенсорів:

- 9-ти осьовий інерційний сенсор – робить цю плату ідеальною для реалізації носимих пристроїв;
- сенсор вологості і температури – використовується для високоточних вимірювань умов навколишнього середовища;

					КС КРБ 123.232.00.00 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		25

- барометричний сенсор – забезпечує можливість створення простої метеостанції
- мікрофон – використовується для захоплення та аналізу звуку в режимі реального часу;
- сенсор жестів, наближення, кольору та інтенсивності світла – дозволяє оцінювати освітленість приміщення, а також визначати чи наближається хтось до місця, де встановлена плата.

Arduino Nano 33 BLE Sense є еволюцією традиційного Arduino Nano, але має набагато потужніший процесор, nRF52840 від Nordic Semiconductors, 32-розрядний процесор ARM® Cortex®-M4, що працює на частоті 64 МГц. Це дозволить створювати більші програми, ніж з Arduino Uno (він має 1 МБ програмної пам'яті, в 32 рази більше), і з набагато більшою кількістю пам'яті (ОЗП в 128 разів більше).

Основний процесор включає в себе додаткові функції, як-от з'єднання Bluetooth через NFC і режими надзвичайно низького енергоспоживання.

Основною особливістю цієї плати, окрім вражаючого вибору сенсорів, є можливість запускати на ній програми Edge Computing (AI) за допомогою TinyML. Розробники можуть створювати свої моделі машинного навчання за допомогою TensorFlow Lite і завантажувати їх на плату за допомогою Arduino IDE. Основні технічні параметри цього виду мікроконтролерів Arduino показано у таблиці 2.1.

Таблиця 2.1 – Технічні параметри Arduino Nano 33 BLE Sense

Компонент	Характеристика
Мікроконтролер	nRF52840
Робоча живлення	3.3В
Вхідна напруга (обмеження)	21В
Постійний струм на виводах вводу/виводу	15 мА
Тактова частота	64МГц
Об'єм флеш-пам'яті	1МБ (nRF52840)

Компонент	Характеристика
Оперативна пам'ять	256КБ (nRF52840)
Постійна пам'ять	відсутня
Цифрові входи / виходи	14
PWM виводи	Всі цифрові виводи
UART	1
SPI	1
I2C	1
Аналогові входи	8 (ADC 12 bit 200 ks)
Аналогові виходи	Тільки через PWM
Зовнішні переривання	Всі цифрові виводи
LED_BUILTIN	13
USB	Власний nRF52840
IMU	LSM9DS1
Мікрофон	MP34DT05
Жест, світло, близькість	APDS9960
Барометричний тиск	LPS22HB
Температура, вологість	HTS221

На рисунку 2.2 показано візуалізовано організацію Nano 33 BLE Sense в контексті схеми електричної принципової.

					КС КРБ 123.232.00.00 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		27

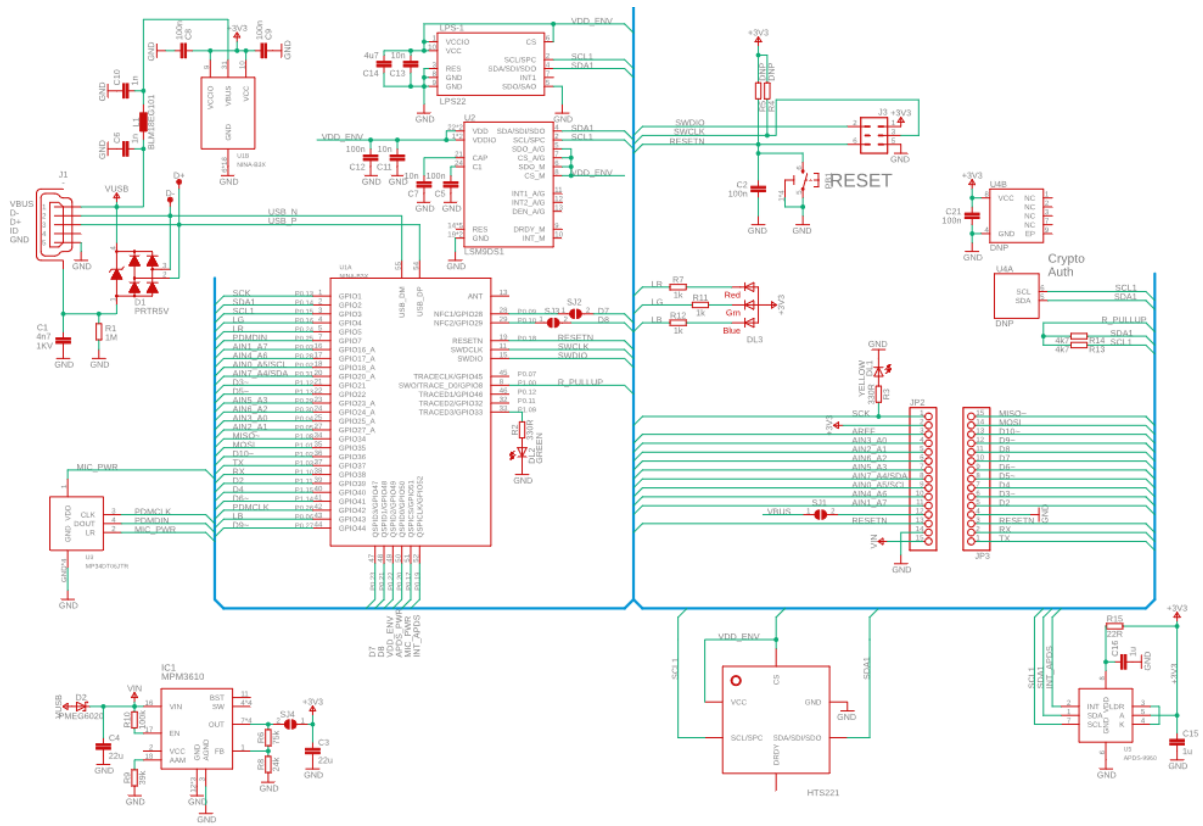


Рисунок 2.2 – Організація схеми електричної принципової Nano 33 BLE Sense

Організація пам'яті та регістрів представлено на рисунку 2.3. Процесор nRF52840 містить 1 МБ флеш-пам'яті та 256 КБ оперативної пам'яті, які можна використовувати для зберігання коду та даних.

Процесор і периферійні пристрої з EasyDMA можуть отримати доступ до пам'яті через багаторівневе з'єднання АНВ. ЦП також може отримати доступ до периферійних пристроїв через багаторівневе з'єднання АНВ.

Змн.	Арк.	№ докум.	Підпис	Дата

КС КРБ 123.232.00.00 ПЗ

Арк.

28

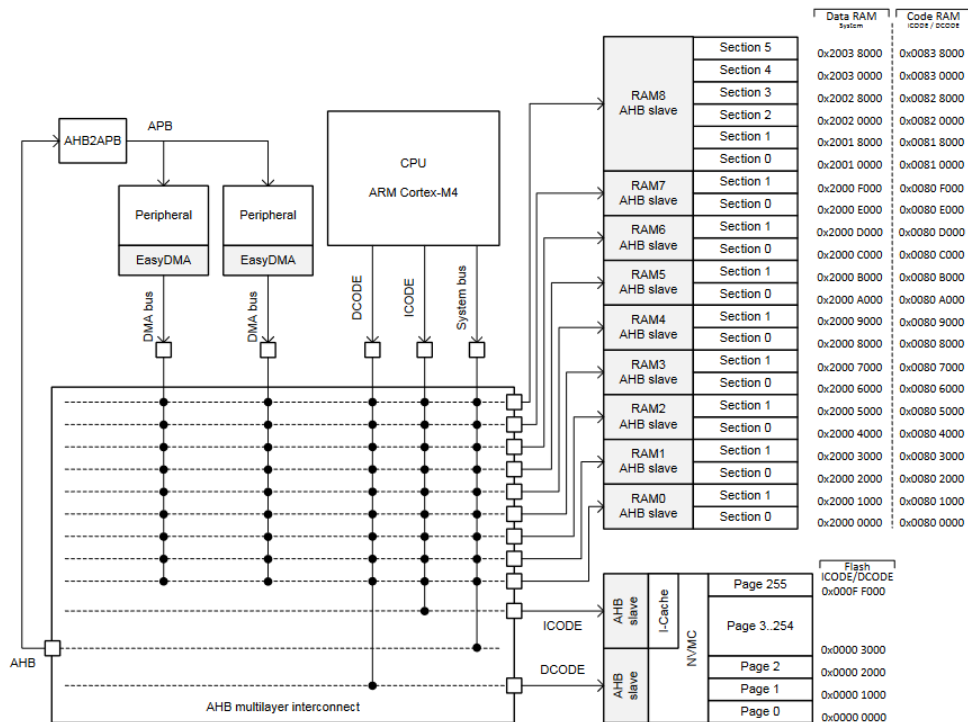


Рисунок 2.3 – Організація пам'яті процесора nRF52840

Джерело живлення складається з ряду LDO і DC/DC стабілізаторів, які використовуються для максимізації енергоефективності системи.

Цей пристрій має такі характеристики джерела живлення:

- вбудовані регулятори LDO та DC/DC;
- режими ввімкнення/вимкнення глобальної системи;
- індивідуальне управління живленням секції RAM для всіх режимів системи;
- пробудження аналогового або цифрового виводу у вимкненій системі;
- обладнання диспетчера для керування скиданням увімкнення живлення, відключенням та збоєм живлення;
- автоматично керовані режими оновлення для регуляторів LDO та DC/DC для максимальної ефективності;
- зовнішнє живлення схем;
- окреме джерело USB.

Загальна структура та організація виводів Nano 33 BLE Sense показана на рисунку 2.4.

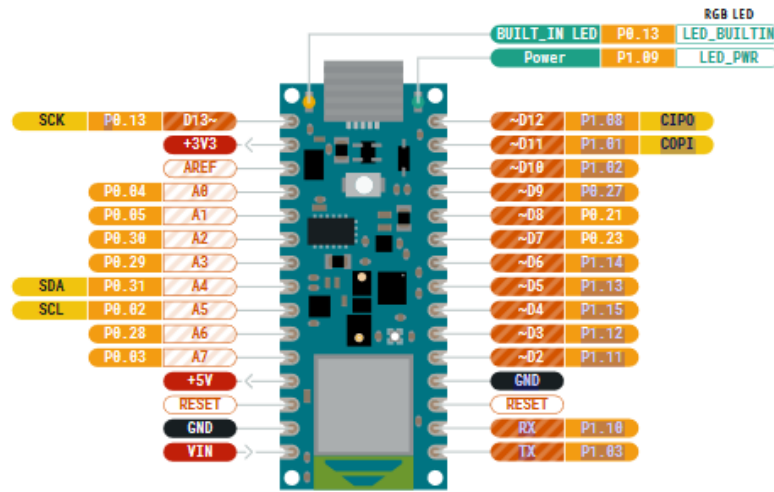


Рисунок 2.4 – Загальна структура та організація виводів використовуваного Arduino

Більш детальна організація Nano 33 BLE з виводами і сенсорами показана на рисунку 2.5.

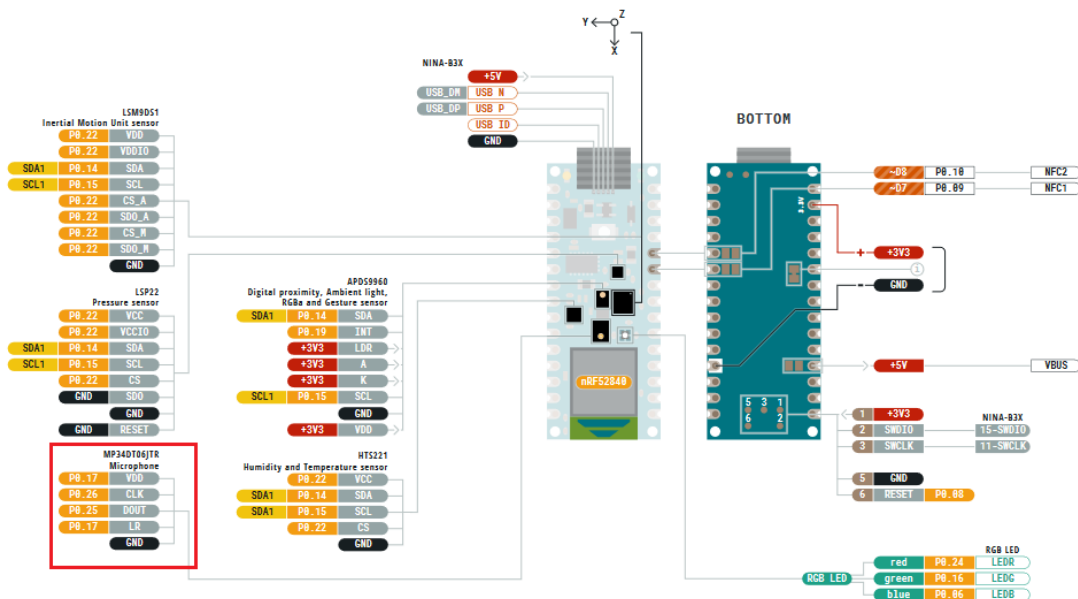


Рисунок 2.5 – Детальна організація виводів і сенсорів Arduino Nano 33

При організації комп'ютерної системи розпізнавання та класифікації звукових сигналів важливу роль відіграє мікрофон, який на рисунку 2.5 виділено червоною рамкою. Тому доцільно розглянуто його технічні характеристики та особливості застосування для того, щоб мати уявлення про обмеження та якість прийому звукового сигналу.

2.3 Характеристики сенсора-мікрофона MP34DT05-A

Arduino Nano 33 має вбудований сенсор мікрофону MP34DT05-A, що представляє собою ультракомпактний, малопотужний, всеспрямований цифровий MEMS-мікрофон із ємнісним чутливим елементом та інтерфейсом ІС.

Чутливий елемент, здатний виявляти акустичні хвилі і виготовляється за допомогою спеціального процесу мікрообробки кремнію, призначеного для виробництва аудіо сенсорів.

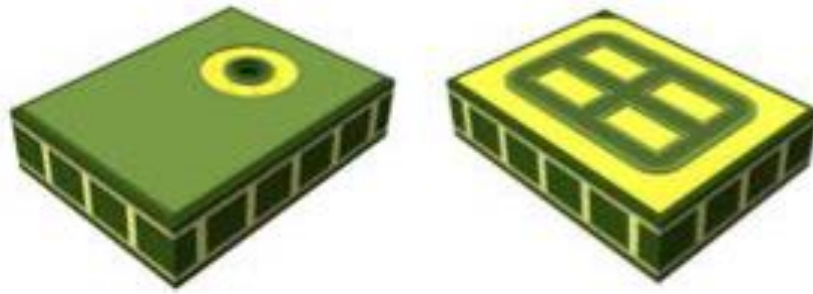
Інтерфейс ІС виготовляється з використанням CMOS, що дозволяє розробити спеціальну схему, здатну надавати цифровий сигнал ззовні у форматі PDM.

MP34DT05-A – цифровий мікрофон з низьким рівнем спотворень із співвідношенням сигнал/шум 64 дБ і чутливістю $-26 \text{ DBFS} \pm 3 \text{ дБ}$. MP34DT05-A доступний у сумісному з SMD, захищеному від електромагнітних шумів корпусі та гарантовано працюватиме в розширеному діапазоні температур від $-40 \text{ }^\circ\text{C}$ до $+85 \text{ }^\circ\text{C}$. На рисунку 2.6 показано зовнішній вигляд MP34DT05-A.

Мікрофон, такого типу, який показаний на рисунку 2.6 може застосовуватися у різних сферах, зокрема при реалізації комп'ютерних систем у вигляді:

- мобільних терміналів;
- ноутбуків та ультрабуків;
- портативних медіаплеєрів;
- системах VoIP;
- системах розпізнавання мови;
- пристроях електронного навчання A/V;
- ігрових систем та пристроїв вводу віртуальної реальності;
- цифрових фото- та відеокамерах;
- протиугінних системах.

					КС КРБ 123.232.00.00 ПЗ	Арк.
						31
Змн.	Арк.	№ докум.	Підпис	Дата		



HCLGA - 4LD (3 x 4 x 1 mm)

Рисунок 2.6 – Вигляд MP34DT05-A

Чутливий елемент MP34DT05-A представляє собою акустичний датчик, що складається з провідної рухомої пластини та нерухомої пластини, розміщених у крихітному кремнієвому чіпі. Цей сенсор перетворює звуковий тиск у зміни пов'язаної ємності між цими двома пластинами.

Напруги, що перевищують зазначені як «абсолютні максимальні показники» у таблиці 2.2, можуть спричинити постійне пошкодження пристрою. Це лише показник напруги, і функціональна робота пристрою в цих умовах не передбачається. Підтримка максимальних умов такого режиму протягом тривалого періоду може вплинути на надійність пристрою.

Таблиця 2.2 – Характеристики і мітки виводів

Позначення	Характеристика	Максимальне значення	Од. вимірювання
V_{dd}	Напруга живлення	-0,3 до 5	В
V_{in}	Вхідна напруга будь-якому керуючого виводу	-0,3 до V_{dd} +0,3	В
T_{STG}	Діапазон температур зберігання	-40 до +125	$^{\circ}C$
ESD	Захист від електростатичного розряду	± 2000 (HBM)	В
		± 200 (MM)	
		± 750 (CBM)	
ESD	Стандарт EN 55024:2010	± 15000 (CBM)	В

Змн.	Арк.	№ докум.	Підпис	Дата

КС КРБ 123.232.00.00 ПЗ

Арк.

32

Позначення та розташування виводу мікросхеми мікрофона показано на рисунку 2.7, а призначення кожного виводу проаналізовано у таблиці 2.3.

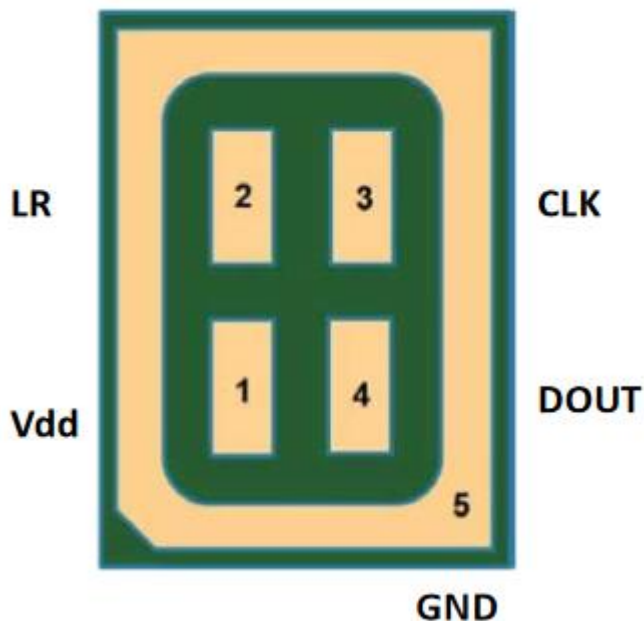


Рисунок 2.7 – Виводи MP34DT05-A

Таблиця 2.3 – Характеристики і мітки виводів

№ виводу	Мітка виводу	Призначення
1	V _{dd}	Живлення
2	LR	Вибір лівого або правого каналу
3	CLK	Синхронізація входу
4	DOUT	Лівий/Правий канал виходу даних PDM
5	GND	Заземлення

Геометричні розміри в одиницях вимірювання міліметри та розташування основних компонентів MP34DT05-A представлено на рисунку 2.8.

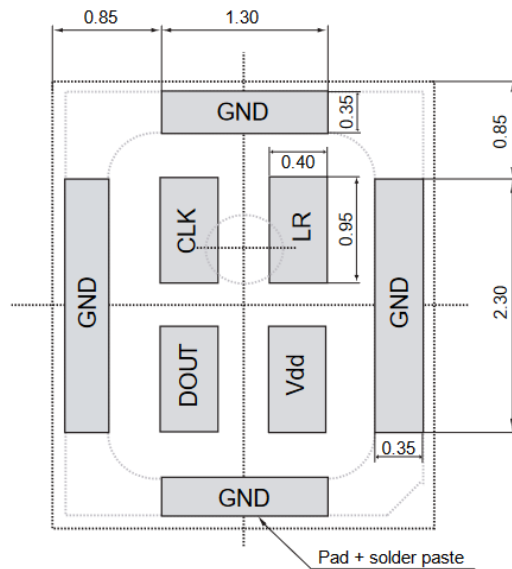


Рисунок 2.8 – Геометричні розміри (мм) та розташування компонентів мікрофона

Таким чином, застосування Arduino Nano 33 для побудови комп'ютерної системи розпізнавання та аналізу звукових сигналів є як технічно, так і економічно обґрунтованим. Наступний крок полягає в аналізі даних, які інтерпретують звукові сигнали, зокрема їхній спектр та реалізації.

2.4 Аналіз природи звукових сигналів

Будь-який об'єкт, який вібрує, виробляє звукові хвилі, наприклад, камертон чи навіть мобільний телефон. Вони видають характерний звук при вібрації. Утворені вібрації формують звук. Тепер, коли об'єкт вібрує, молекули повітря коливаються в різні сторони з положення спокою об'єкта і передають свою енергію сусіднім молекулам. Це призводить до передачі енергії від однієї молекули до іншої, яка, в свою чергу, створює звукову хвилю. На рисунку 2.9 наведено формування звукової хвилі камертоном.

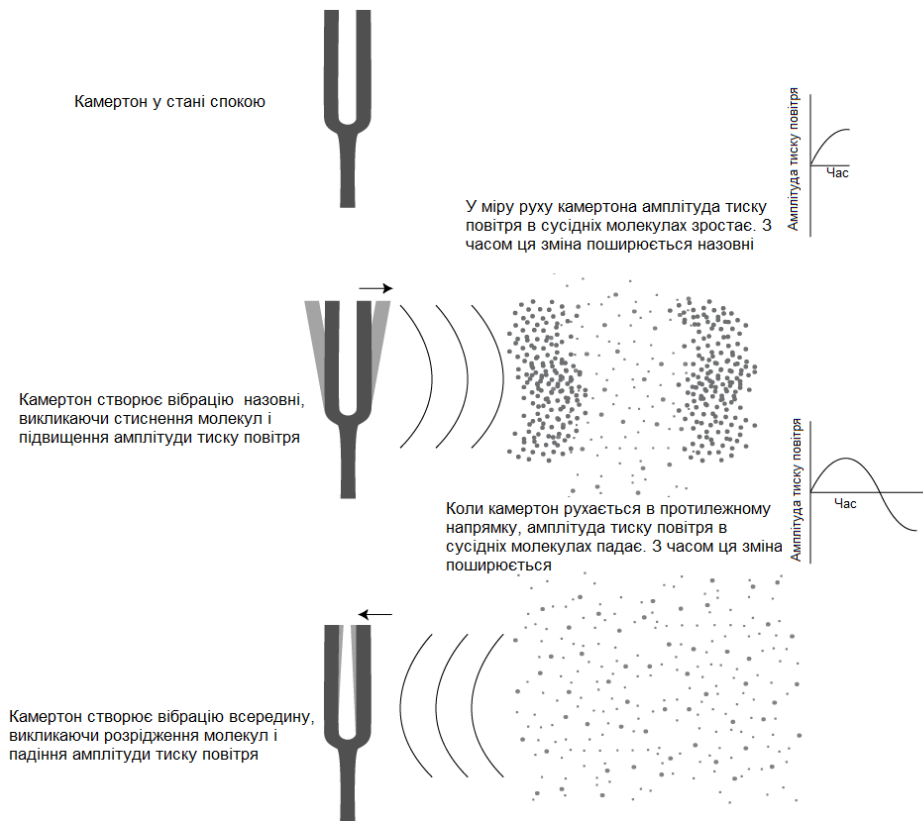


Рисунок 2.9 – Формування звукової хвилі камертоном

Таке представлення звуку (рисунок 2.9) відоме як звуковий сигнал. Аудіосигнал має наступні кілька важливих параметрів:

- довжину хвилі;
- період;
- амплітуду та частоту.

Ці параметри дуже важливі і корисні для побудови моделі та вирішення різних бізнес-задач

Амплітуда – це максимальне зміщення або відхилення частинки від положення спокою. Іноді положення спокою також відоме як середнє положення. На зображенні нижче (рисунок 2.10) чорна горизонтальна лінія представляє положення рівноваги (або середнє, або спокою), а максимальне відхилення від цієї чорної лінії – амплітуда.

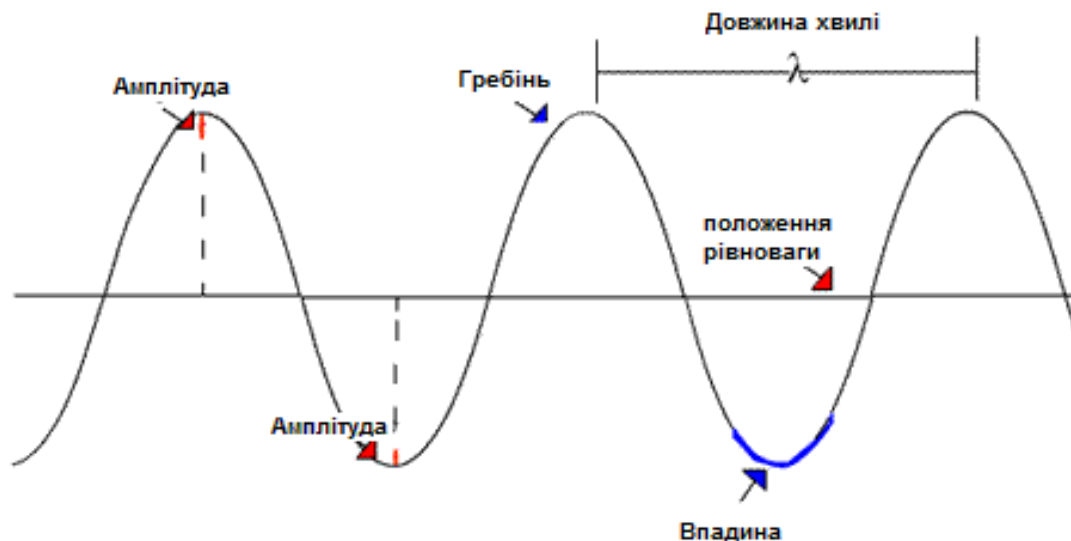


Рисунок 2.10 – Представлення амплітуди звукового сигналу

Іншими словами, амплітуда вказує на те, наскільки частинка відхиляється від нормального непорушеного положення. Це висота гребеня або глибина впадини, як показано пунктирними лініями на зображенні вище. Перейдемо до аналізу того, як поширюється звук.

Аудіохвилі поширюються у вигляді періодів або циклів. Період – це один повний рух хвилі вгору і вниз до її середнього положення, як проілюстровано на рисунку 2.11.



Рисунок 2.11 – Ілюстрація періоду звукового сигналу

Частота – це показник, що виражає наскільки швидко змінюється звуковий сигнал з часом. Його також можна назвати кількістю періодів або циклів в

секунду. Іншими словами, кількість повторень хвилі (повторна хвиля дорівнює циклу) за одиницю часу називається частотою. Візуальне представлення частоти сигналу показано на рисунку 2.12.



Рисунок 2.12 – Візуалізація частоти сигналу

Варто відзначити, що період може початися в будь-якому місці звукової хвилі. Однак саме період звукової хвилі визначає частоту, а не те, з чого почався цикл. Коли частота збільшується вдвічі, то кількість хвиль також збільшується вдвічі в будь-який момент часу.

Низькі частоти матимуть довгохвильові або більші періоди, а високі частоти матимуть коротшу довжину хвилі та коротші періоди.

До цього часу проаналізовано поняття аудіосигналу, як він поширюється та змінюється у часі. Тепер потрібно розглянути у яких форматах зберігаються аудіофайли. Існує два різних типи аудіосигналів: аналогові та цифрові.

Аналоговий сигнал – це безперервна хвиля, яка змінюється протягом певного періоду часу. Між будь-якими двома послідовними екземплярами часу існує нескінченна кількість вибірок. На рисунку 2.13 показано приклад аналогового і цифрового сигналів, а також реального звукового сигналу.

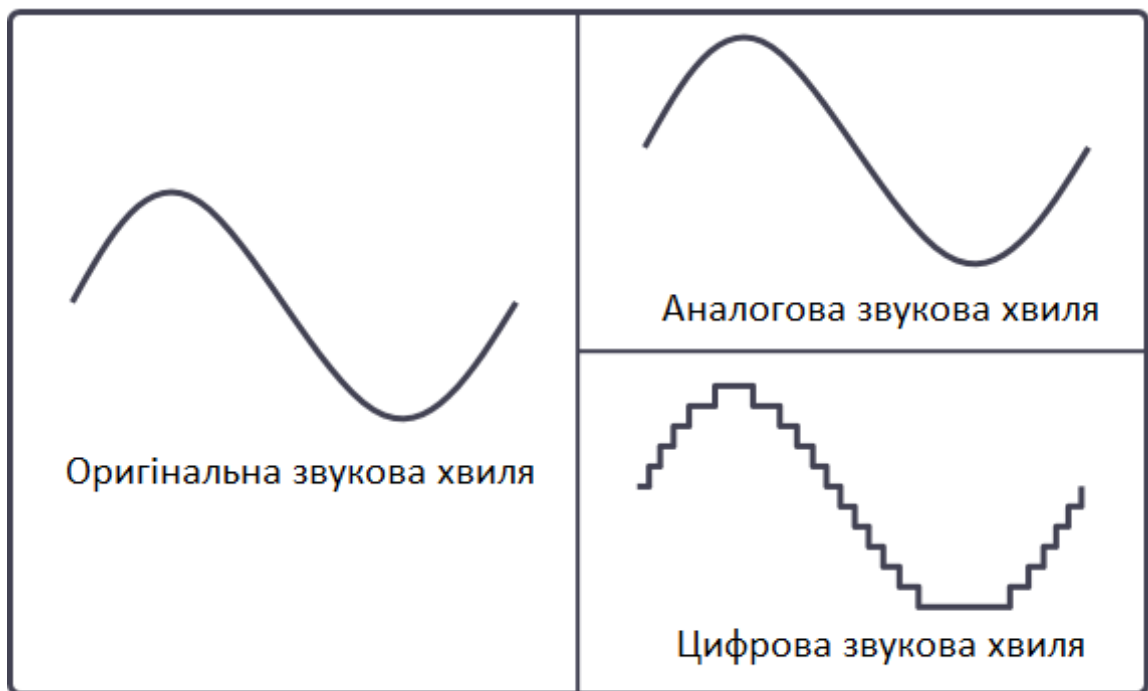


Рисунок 2.13 – Представлення різних типів звукових сигналів

З іншого боку, як видно із зображення вище, цифрові сигнали є дискретним представленням сигналу за певний період часу.

Тут між будь-якими двома послідовними екземплярами часу існує лише кінцева кількість значень. Сигнали можуть мати лише одне значення в будь-який момент часу: або 0, або 1, тобто є двійковими.

Також можна сказати, що це дискретна хвиля, яка несе інформацію в двійковій формі. Для того, щоб працювати над задачею із застосуванням інформаційних технологій, пов'язаною з аудіоданими та обробкою аудіосигналів, потрібно виконувати перетворення аналогових сигналів у цифрові.

Це можна зробити через дискретизацію. Дискретизація відома як перетворення аналогових сигналів у цифрові. На зображенні нижче (рисунок 2.14) зелена хвиля є аналоговим сигналом, а червоні точки, позначені на ній, є дискретними значеннями цієї кривої.

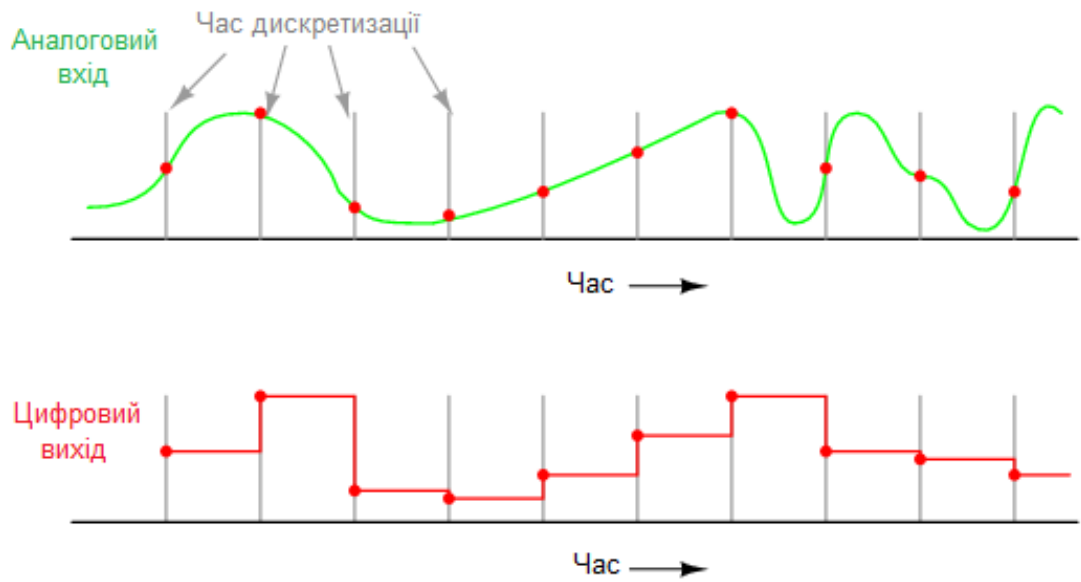


Рисунок 2.14 – Перетворення аналогового сигналу у цифровий

Під час вибірки після фіксованого інтервалу часу вимірюється значення аналогового сигналу, і цей вибірковий сигнал пізніше може бути використаний для відновлення вихідного безперервного сигналу. Тепер середня кількість вибірок, захоплених за одну секунду, відома як частота дискретизації. Приклад різної низької і високої частоти дискретизації показано на рисунку 2.15.

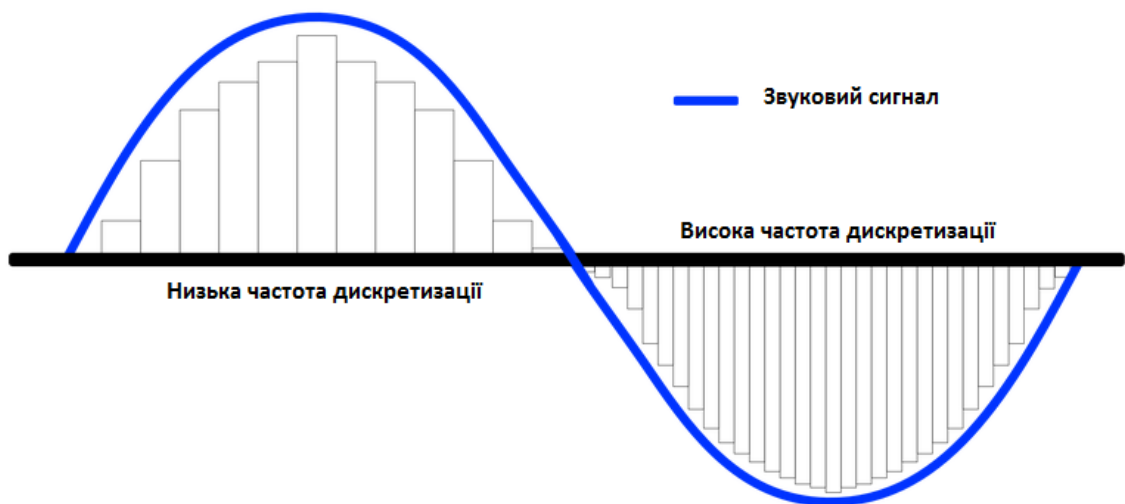


Рисунок 2.15 – Низька і висока частота дискретизації

Змн.	Арк.	№ докум.	Підпис	Дата

Чим вища частота дискретизації, тим вищою буде якість звуку і також запис буде займати більше місця в пам'яті.

Існує два способи представлення звукового сигналу через часову область та спектрограму.

У часовій області звук представлений амплітудою, як функцією часу, тобто амплітуда записується через різні проміжки часу. Простіше кажучи, це графік залежності між амплітудою і часом: вісь X – час, а вісь вісь Y – амплітуда. Залежність амплітуди від час показана на рисунку 2.16.

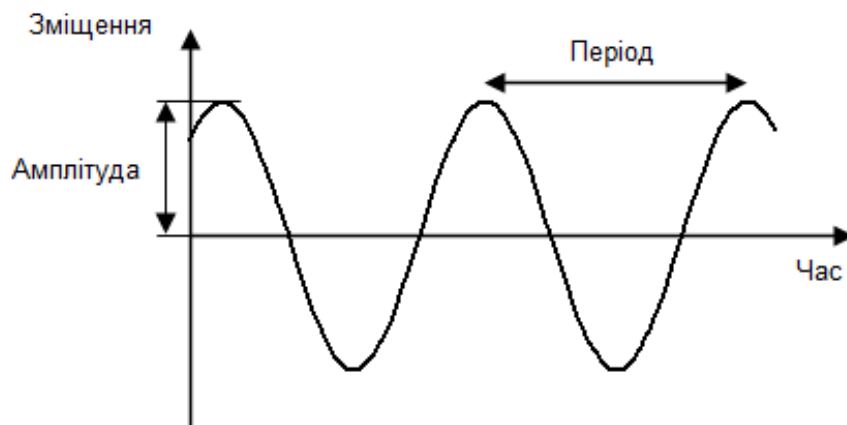


Рисунок 2.16 – Звуковий сигнал в часовій області

Спектрограма (рисунок 2.17) є ще одним представленням звукового сигналу, що представляє собою двовимірний (2D) графік залежності між часом і частотою і має третій вимір, який представляється кольором.

Кожне значення спектрограми представляє амплітуду частоти в конкретний момент часу в термінах інтенсивності кольорів. Це простіше кажучи – спектр частот, який змінюється з часом, де вісь X – це час, де час йде зліва (найстарший) направо (наймолодший). Вісь Y – це частота, де найнижчі частоти знаходяться внизу, а найвищі – угорі. Значення – це амплітуди певної частоти у певний момент часу.

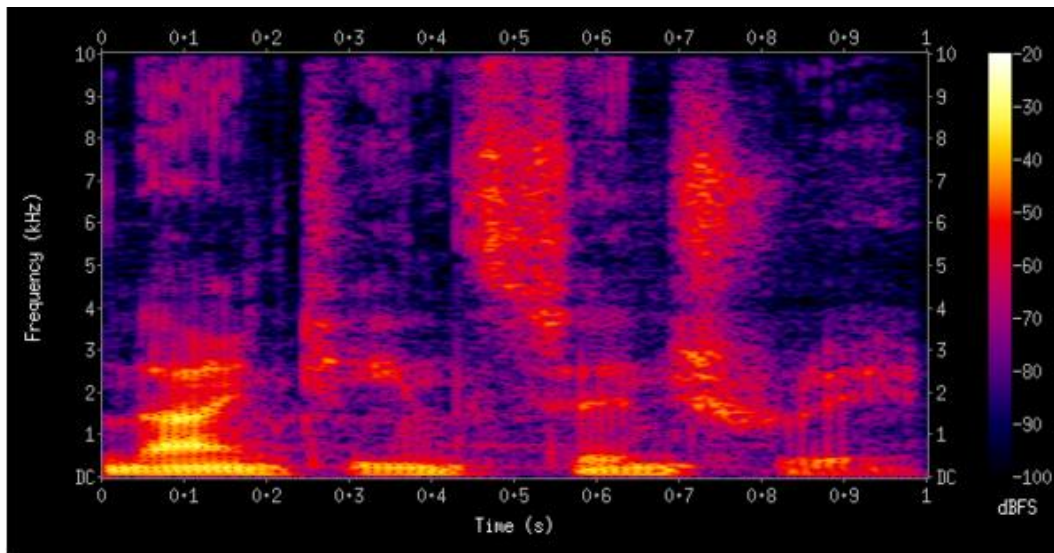


Рисунок 2.17 – Спектрограма звукового сигналу

Спектрограма є одним із видів теплової карти, де інтенсивність зображується різними кольорами та яскравістю. Нижній колір на спектрограмі представляє нижчу інтенсивність і нижчі амплітуди, а темніші кольори відповідають прогресивно сильнішим (або гучнішим) амплітудам.

Спектрограму також можна розглядати як візуальне уявлення сили сигналу або «гучності» сигналу у часі на різних частотах, представлених у певній формі хвилі.

Тепер маючи обґрунтований вибір апаратного забезпечення та проаналізовані параметри звукових сигналів, перейдемо до реалізації програмного забезпечення розпізнавання і класифікації звукових сигналів на прикладі бінарної класифікації звукових сигналів спеціальних транспортних засобів та звичайних.

РОЗДІЛ 3 РЕАЛІЗАЦІЯ ПРОГРАМНОЇ МОДЕЛІ РОЗПІЗНАВАННЯ І КЛАСИФІКАЦІЇ ЗВУКОВИХ СИГНАЛІВ

3.1 Побудова процедури реалізації програмного забезпечення аналізу звуків сигналів і препроцесинг даних

За даними Національного бюро реєстрації злочинів в США, майже 24012 людей помирають щодня через запізнилу медичну допомогу. Багато постраждалих чекають допомоги на місці, а затримка відповідних служб коштує їм життя.

Причинами цього можуть бути різні фактори: від машин швидкої допомоги, які застрягли в заторі, до того, що пожежна команда не може вчасно приїхати на місце через корки.

Розв'язання вищеприписаної проблеми полягає у створенні системи, яка може автоматично виявляти автомобілі спеціальних та екстрених служб до того, як вони досягнуть світлофора, і відповідно зміниться сигнал світлофора.

Експеримент пропонується виконати на наборі даних, який містить два файли: перший файл з розширенням *.wav містить звуки аварійних автомобілів тривалістю приблизно 23 хвилини, а інший файл – звуки не аварійних транспортних засобів тривалістю 27 хвилин.

Для того, щоб досягти поставленої мети пропонується використати мову програмування Python та виконати наступні кроки:

- імпортувати необхідні бібліотеки;
- виконати читання аудіоданих;
- дослідити та виконати препроцесинг даних;

					КС КРБ 123.232.00.00 ПЗ			
<i>Змн.</i>	<i>Арк.</i>	<i>№ докум.</i>	<i>Підпис</i>	<i>Дата</i>				
<i>Розроб.</i>		Сахарчук В.В.			<i>Реалізація програмної моделі розпізнавання і класифікації звуків сигналів</i>	<i>Літ.</i>	<i>Арк.</i>	<i>Аркуші</i>
<i>Перевір.</i>		Яцишин В.В.					42	
<i>Реценз.</i>						<i>ТНТУ, каф. КС, гр. СІс-43</i>		
<i>Н. Контр.</i>		Луцик Н.С.						
<i>Затверд.</i>		Осухівська Г.М.						

- декомпозиція вибірки даних на навчальну і тестову;
- побудова моделі машинного навчання;
- навчання і тестування моделі;
- формування рішення щодо приналежності звуку до відповідного класу.

Послідовність виконання процесів щодо класифікації аудіоданих показано на рисунку 3.1

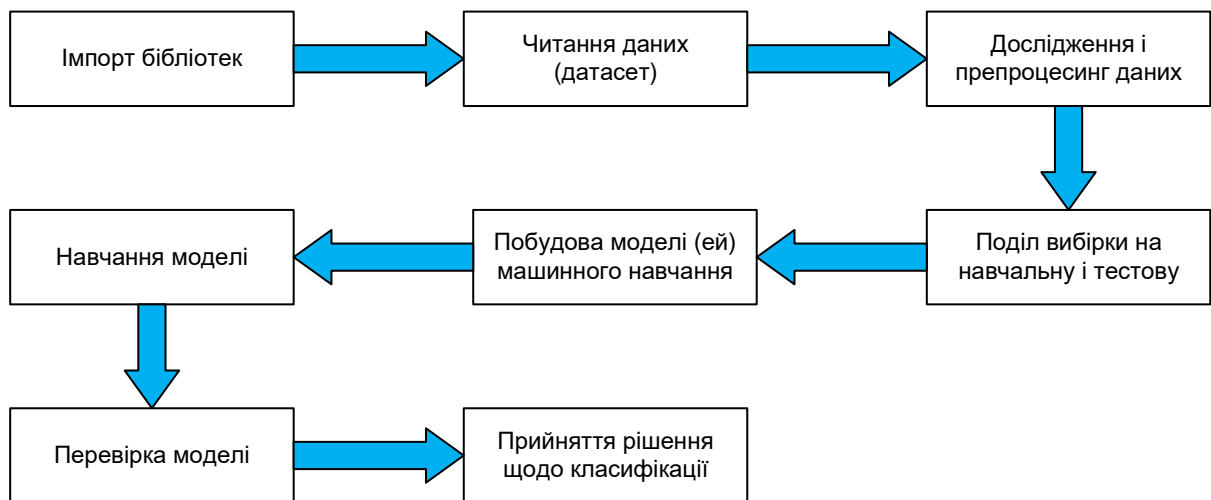


Рисунок 3.1 – Алгоритм реалізації програмного забезпечення класифікації звукових сигналів

Бібліотеки, що використовуються для класифікації аудіо

- Librosa – використовується для попереднього опрацювання звуку, аналізу аудіо та музики;
- Scipy- бібліотека, що використовується для науково-технічних обчислень і містить модулі для опрацювання сигналів, обробки зображень і лінійної алгебри;
- Numpya – застосовується для роботи з масивами;
- IPython.display – використовується для відтворення аудіо;
- Matplotlib – використовується для створення двовимірних графіків і візуалізації даних;

– Sklearn – бібліотека побудована на NumPy і SciPy, що містить інструменти для прогнозного аналізу даних і буде використовуватися для поділу даних на train та test;

– Keras – бібліотека, що володіє інструментами побудови нейронних мереж, імпортування відповідних шарів та реалізації нейронних мереж.

Після імпорту необхідних бібліотек потрібно завантажити аудіодані (за допомогою функції завантаження) з бібліотеки librosa, як показано у лістингу 3.1.

Лістинг 3.1 – Імпорт бібліотек та читання аудіоданих

```
# importing emergency file:
emergency, sample_rate = lib.load('audio/emergency.wav',
sr=16000)

# importing non-emergency file:
non_emergency, sample_rate = lib.load('audio/non
emergency.wav', sr=16000)
```

Після цього доцільно знайти тривалість аудіозаписів за допомогою функції `get_duration()` з бібліотеки librosa (лістинг 3.2)

Лістинг 3.2 – Пошук тривалості аудіозаписів

```
len1 = lib.get_duration(emergency, sr=16000)
len2 = lib.get_duration(non_emergency, sr=16000)
print('The duration of the Emergency clip is {} mins and Non-
Emergency clip is {} mins.'.format(round(len1/60,2),
round(len2/60,2)))
```

Потрібно пам'ятати, що хвилі звукового сигналу аналогові і для їх перетворення у цифрові, тобто двійкові та дискретні сигнали, застосовується процес дискретизації. Таким чином, частота дискретизації (змінна «sr»=16000) зчитує два вищезазначені аудіозаписи. Ця частота дискретизації означає, що 16000 — це середня кількість вибірок, записаних за секунду.

					КС КРБ 123.232.00.00 ПЗ	Арк.
						44
Змн.	Арк.	№ докум.	Підпис	Дата		

Після завантаження даних, наступним кроком є підготовка аудіоданих. Оскільки аудіофайли не можуть бути використані безпосередньо в алгоритмах, цей крок підготовки аудіоданих є дуже важливим. Розіб'ємо дані аудіозапису на послідовності аудіо, щоб навчити модель глибокого навчання виконувати класифікацію аудіо.

Процедура підготовки звукового сигналу полягає в наступному. Скажімо, у нас є блок із дев'яти клітинок, як показано нижче (рисунок 3.2). При цьому кожна клітинка представляє одну секунду аудіо. Отже, маємо дев'ять секунд звукових даних.

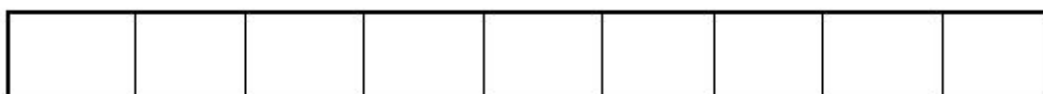


Рисунок 3.2 – Представлення 9 с звукових даних

Далі розіб'ємо цю послідовність тривалістю дев'ять секунд на аудіозаписи по дві секунди кожна, як показано на рисунку 3.3.

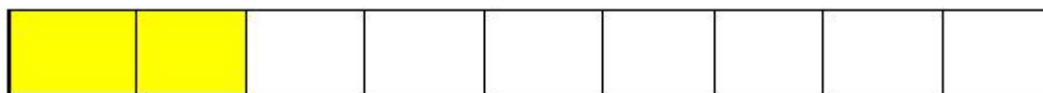


Рисунок 3.3 – Розбиття аудіозаписів тривалістю по 2 с

Таким чином, одержано перший фрагмент аудіо (розміром 2 секунди), а щоб витягти інші фрагменти потрібно переміщати 2-секундне вікно з кроком 1. Процедуру переміщення (ковзання) вікна показано на рисунку 3.4.

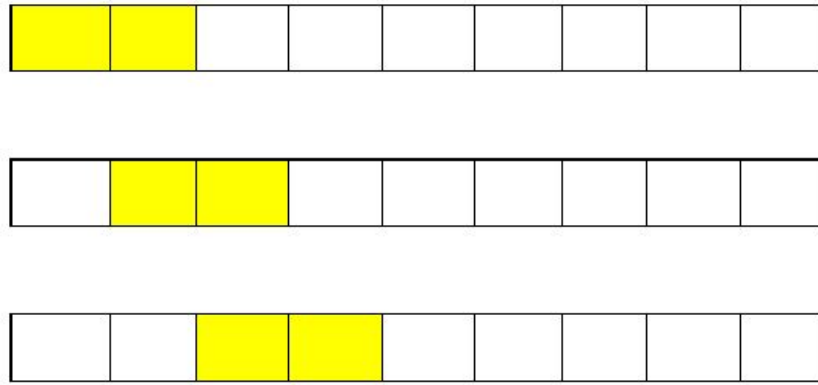


Рисунок 3.4 – Ковзання вікна по аудіозапису

Отже, у результаті проведення процедури, показаної на рисунку 3.4, з дев'яти секунд послідовності аудіоданих можна отримати вісім аудіо фрагментів або послідовностей, кожна з яких має довжину 2 секунди. У лістингу 3.3 наведено функцію розбиття аудіофайлу на послідовності.

Лістинг 3.3 – Функція розбиття аудіозапису на послідовності

```
# User-defined Function for Audio chunks audio data is the
array
def audio_chunks(audio_data, num_of_samples=32000, sr=16000):
    # empty list to store new audio chunks formed
    data=[]
    for i in range(0, len(audio_data), sr):
        # creating the audio chunk by starting with the first
second & sliding the 2-second window one step at a time
        chunk = audio_data[i: i+ num_of_samples]
        if(len(chunk)==32000):
            data.append(chunk)
    return data
```

Тепер частота дискретизації для кожного аудіофайлу становить 16 000, тож аудіо тривалістю 2 секунди з частотою дискретизації 16 000 буде мати 32 000 вибірок. Після виклику функції (лістинг 3.4) одержуємо кількість блоків зі звуковими сигналами машин екстрених служб в кількості 1374 і неаварійних 1628.

Лістинг 3.4 – Виклик функції розбиття звукового сигналу

```
# Calling the above function audio_chunks to create #seperate
chunks for both Emergency and non-emergency #vehicles:
emergency = audio_chunks(emergency)
non_emergency = audio_chunks(non_emergency)
print('The number of chunks of Emergency is {} and Non-
Emergency is {}'.format(len(emergency), len(non_emergency)))
```

Наступний крок полягає у дослідженні того, як виглядає графік залежності амплітуди від часу. Для цього потрібно спочатку потрібно реалізувати скрипт, який повертає аудіокліп, щоб можна було відтворити кожен з екстрених та неаварійних звуків (лістингу 3.5).

Лістинг 3.5 – Відтворення звукових сигналів екстрених служб

```
# Emergency Sound:
ipd.Audio(emergency[45], rate=16000)
# Non-Emergency Sound:
ipd.Audio(non_emergency[29], rate=16000)
```

Для візуалізації спектру звукових сигналів необхідно реалізувати програмний код, представлений у лістингу 3.6.

Лістинг 3.6 – Візуалізація залежностей амплітуди від часу

```
# Emergency Sound
plt.figure(figsize=(14,4))
plt.plot(np.linspace(0, 2, num=32000),emergency[32])
plt.title('Emergency Sound')
plt.xlabel('Time')
plt.ylabel('Amplitude')
plt.show()
# Non-Emergency Sound
plt.figure(figsize=(14,4))
plt.plot(np.linspace(0, 2, num=32000),non_emergency[33])
plt.title('Non-Emergency Sound')
```

					КС КРБ 123.232.00.00 ПЗ	Арк.
						47
Змн.	Арк.	№ докум.	Підпис	Дата		

```
plt.xlabel('Time')
plt.ylabel('Amplitude')
plt.show()
```

У результаті вищенаведеного лістингу одержують графіки, показані на рисунку 3.5.

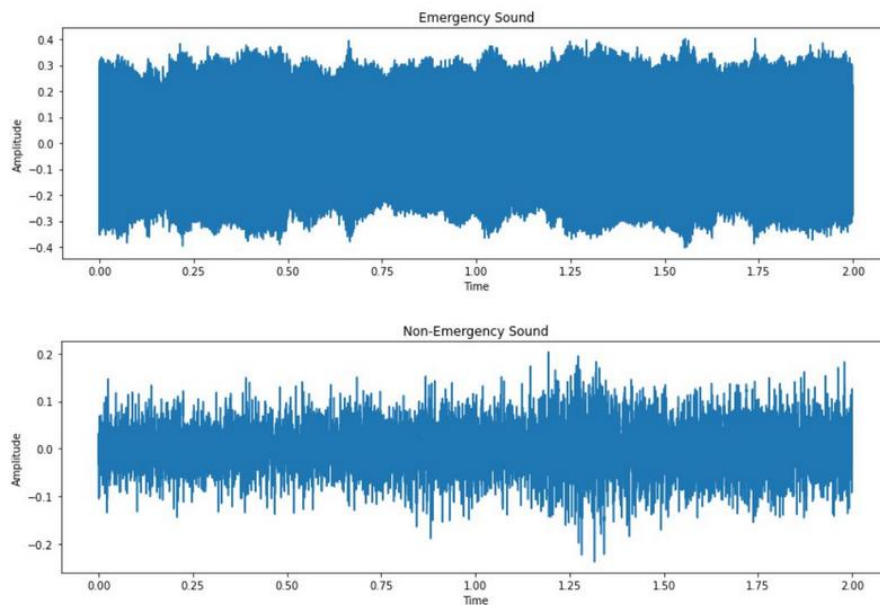


Рисунок 3.5 – Візуалізація звукових сигналів автомобілів екстрених служб та звичайних авто

Тепер із створених вище аудіо фрагментів або послідовностей підготуємо дані для навчання та тестування майбутніх моделей. Першим кроком є поєднання звукових сигналів екстрених і звичайних авто. У даному випадку прийнято позначити звуки екстрених авто з міткою 0, а звичайні як мітку 1. Отже потрібно розв'язати задачу бінарної класифікації. Лістинг 3.7 виконує нанесення міток на послідовності звукових сигналів.

Лістинг 3.7 – Формування міток для класів звуків

```
# Step 1: Combining the Emergency and Non Emergency audio
chunks
audio = np.concatenate([emergency, non_emergency])
```

					КС КРБ 123.232.00.00 ПЗ	Арк.
						48
Змн.	Арк.	№ докум.	Підпис	Дата		


```

# Step 2: Assigning labels
labels1 = np.zeros(len(emergency))
labels2 = np.ones(len(non_emergency))
# concatenate labels
labels = np.concatenate([labels1,labels2])
print('The shape of the combined audio data is {}'.
      .format(audio.shape))
# Train-test splitting:
X_train, X_test, Y_train, Y_test =
train_test_split(np.array(audio), np.array(labels),
stratify=labels, test_size=0.20,

random_state=12, shuffle=True)
print('X_train', X_train.shape)
print('X_test', X_test.shape)
print('')
print('Y_train', Y_train.shape)
print('Y_test', Y_test.shape)

```

У результаті одержано 3002 аудіо послідовності з 32000 вимірами. Тепер модель послідовності в Python очікує три введення:

- кількість реалізацій;
- крок дискретизації;
- довжина ознак.

Виходячи з цього, вхідні дані мають бути тривимірним масивом, однак зараз вони є двовимірним масивом. Тому потрібно змінити форму вхідних даних з урахуванням вимог. Оскільки довжина об'єктів тут відсутня, то можна встановити третій вимір довжину як 1. У лістингу 3.8 наведено приведення двовимірного масиву у тривимірний.

Лістинг 3.8 – Перетворення 2D у 3D

```

# Reshaping the 2-Dimensional array into 3-Dimensional array
by setting the third dimension to 1:
X_train_features = X_train.reshape(len(X_train),-1,1)

```

					КС КРБ 123.232.00.00 ПЗ	Арк.
						49
Змн.	Арк.	№ докум.	Підпис	Дата		

```

X_test_features = X_test.reshape(len(X_test), -1,1)
print('The reshaped X_train array has size:',
      X_train_features.shape)
print('The reshaped X_test array has size:',
      X_test_features.shape)

```

Змінений масив `X_train` має розмір: (2401, 32000, 1), а масив `X_test` - (601, 32000, 1) До цього моменту кроки були однаковими, незалежно від того, використовувалась часова область чи спектрограма.

Таким чином виконано підготовку даних та створено набори навчальної і тестових вибірок. Наступний крок полягає у реалізації двох підходів із застосуванням різних моделей для вирішення задачі бінарної класифікації.

Неопрацьований звуковий сигнал доступний у часовій області, тому можна виконати його безпосередній препроцесинг та почати побудову модель класифікації звуків авто екстрених служб та звичайних. Для початку побудуємо модель за допомогою CNN і LSTM.

3.2 Побудова моделі машинного навчання на основі ознак часової області

Працюючи із звуковими або аудіоданими пропонується застосувати архітектуру нейронної мережі Conv1D, оскільки він використовується для вхідних сигналів, подібних до голосових.

Така модель складається з двох наборів згорткових шарів (convolution) та шарів об'єднання (pooling). Визначення функції, що містить архітектуру моделі, компілятора, а також контрольної точки моделі наведено у лістингу 3.9.

Лістинг 3.9 – Функція побудови моделі нейронної мережі

```

def cnn_model(X_tr):
    inputs = Input(shape= (X_tr.shape[1], X_tr.shape[2]))
    # first Conv1D Layer with 8 filters of height 13:
    conv = Conv1D(8,13, padding='same',
                  activation='relu')(inputs)

```

					КС КРБ 123.232.00.00 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		50

```

conv = Dropout(0.3)(conv)
conv = MaxPooling1D(2)(conv)
# 2nd Conv1D Layer with 16 filters of height 11:
conv = Conv1D(16,11, padding='same',
activation='relu')(inputs)
conv = Dropout(0.3)(conv)
conv = MaxPooling1D(2)(conv)
# Global MaxPooling 1D
conv = GlobalMaxPool1D()(conv)
# Dense Layer
conv = Dense(16, activation='relu')(conv)
outputs = Dense(1,activation='sigmoid')(conv)
model = Model(inputs, outputs)
# Model Compiler:
model.compile(loss='binary_crossentropy',
optimizer='adam', metrics=['acc'])
# Model Checkpoint
model_checkpoint = ModelCheckpoint('best_model_cnn.hdf5',
monitor='val_loss', verbose=1, save_best_only=True,
mode='min')
return model, model_checkpoint

```

Функція у лістингу 3.9 повертає два об'єкти: модель і збережені ваги моделі. Оскільки у роботі розв'язується задача бінарної класифікації звукових сигналів, то доцільно використовувати функцію втрат – `binary_crossentropy`. Виклик функції створення моделі CNN показано у лістингу 3.10.

Лістинг 3.10 – Виклик функції побудови моделі CNN

```

# Calling the model:
model, model_checkpoint = cnn_model(X_train_features)
# Shape and parameters at each layer
model.summary()

```

					КС КРБ 123.232.00.00 ПЗ	Арк.
						51
Змн.	Арк.	№ докум.	Підпис	Дата		

Структура моделі за кількістю шарів та нейронів у них проілюстровано на рисунку 3.6.

Model: "model_6"

Layer (type)	Output Shape	Param #
input_10 (InputLayer)	(None, 32000, 1)	0
conv1d_16 (Conv1D)	(None, 32000, 16)	192
dropout_16 (Dropout)	(None, 32000, 16)	0
max_pooling1d_16 (MaxPooling)	(None, 16000, 16)	0
global_max_pooling1d_8 (GlobalMaxPooling1D)	(None, 16)	0
dense_11 (Dense)	(None, 16)	272
dense_12 (Dense)	(None, 1)	17
Total params: 481		
Trainable params: 481		
Non-trainable params: 0		

Рисунок 3.6 – Структура побудованої моделі CNN

У лістингу 3.11 показано програмний код для навчання моделі та формування підсумку історії функції втрат.

Лістинг 3.11 – Навчання нейронної мережі

```
# model training
history = model.fit(X_train_features, Y_train ,epochs=10,
                    callbacks=[model_checkpoint], batch_size=32,
                    validation_data=(X_test_features,
                    Y_test))
# accuracy for the model evaluation
# load the best model weights
model.load_weights('best_model_cnn.hdf5')
# summarize history for loss
plt.plot(history.history['loss'])
```

```

plt.plot(history.history['val_loss'])
plt.title('model loss')
plt.ylabel('loss')
plt.xlabel('epoch')
plt.legend(['train', 'validation'], loc='upper left')
plt.show()

```

Результат навчання CNN показано на рисунку 3.7. Даний графік показує залежність функції втрат від епох навчання та похибку результатів класифікації на навчальній і тестовій вибірках.

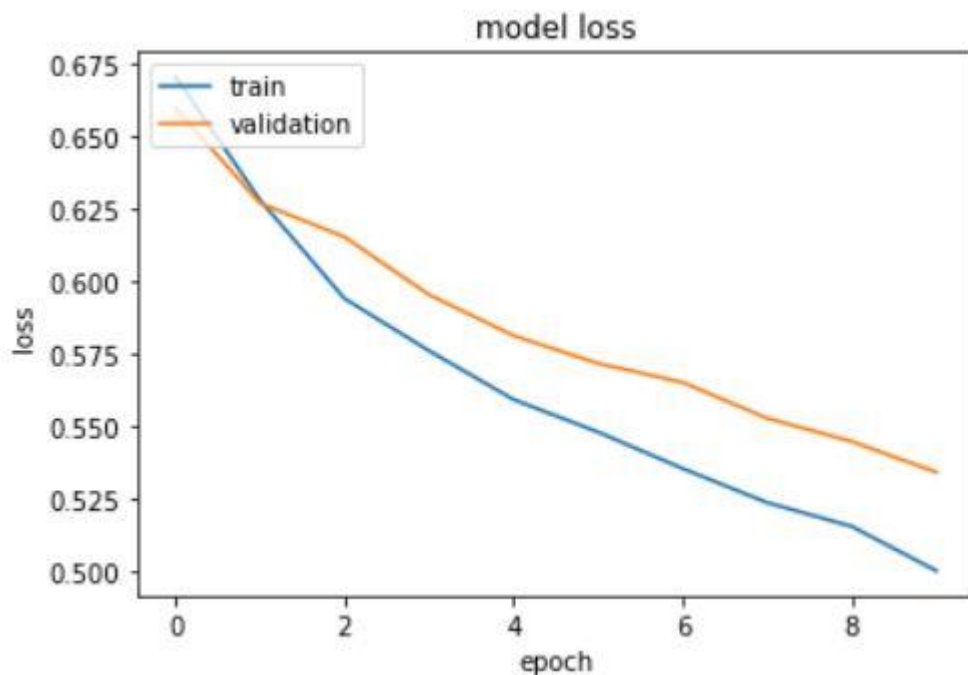


Рисунок 3.7 – Результат навчання мережі при бінарній класифікації

Для оцінювання продуктивності моделі на тестових даних реалізовано скрипт, що приведений у лістингу 3.12.

Лістинг 3.12 – Оцінювання продуктивності моделі

```

# Checking the model's performance on the Test set
_, acc = model.evaluate(X_test_features, Y_test)
print("Validation Accuracy:", acc)

```

У результаті навчання побудована модель забезпечує точність на рівні 77,038% на тестових даних. Тепер, застосуємо цю модель, щоб класифікувати звуки автомобілів екстреної допомоги або сигнали звичайних автомобілів. Для того, щоб прогнозувати і класифікувати такі сигнали використовується програмний код лістингу 3.13.

Лістинг 3.13 – Прогнозування і класифікація звукових сигналів

```
# For Prediction: the input audio
index = 98
test_audio = X_test[index]
# Using IPython.display to play the audio
ipd.Audio(test_audio, rate=16000)
# classification
feature = X_test_features[index]
prob = model.predict(feature.reshape(1,-1,1))
if (prob[0][0] < 0.5):
    pred='emergency'
else:
    pred='non emergency'
print("Prediction:",pred)
```

У результаті виконання лістингу 3.13 одержують аудіофайл, і модель прогнозує його мітку як неаварійний автомобіль. Однак, оскільки точність моделі становить лише 77%, безумовно, можна покращити модель, зокрема, шляхом додавання іншого алгоритму – LSTM. Наступний крок полягає у побудові моделі за допомогою алгоритму LSTM за ознаками часової області.

Форма вхідних послідовностей становить 32000. 32000 семплів означає, що кожен фрагмент аудіо буде мати 32000-часових міток. Це може призвести до великих затрат часу на тренування, а також може виникнути проблема зникаючого градієнтного спуску.

Отже, щоб подолати це, скористаємося простим трюком під час зміни форми фрагментів нижче (лістинг 3.14). Замість того, щоб зберегти довжину об'єктів (третій вимір) як 1, буде використано 160, що зменшить часові мітки з

					КС КРБ 123.232.00.00 ПЗ	Арк.
						54
Змн.	Арк.	№ докум.	Підпис	Дата		

32000 до 200, а потім можна легко використовувати ці змінені послідовності у моделі.

Лістинг 3.14 – Перетворення послідовності звукових сигналів

```
# Reshaping the Audio chunks:
X_train_features = X_train.reshape(len(X_train), -1, 160)
X_test_features = X_test.reshape(len(X_test), -1, 160)
print('The reshaped X_train array has size:',
      X_train_features.shape)
print('The reshaped X_test array has size:',
      X_test_features.shape)
```

У результаті одержимо структуру послідовностей для навчальної вибірки – (2401, 200, 160), а для тестової – (601, 200, 160). Наступний крок полягає у визначенні функції реалізації алгоритму LSTM, який показано у лістингу 3.15.

Лістинг 3.15 – Визначення моделі з використанням LSTM та її виклик

```
# LSTM based deep learning model architecture
def lstm_model(X_tr):
    inputs = Input(shape=(X_tr.shape[1], X_tr.shape[2]))
    # LSTM Layer 1
    x = LSTM(128)(inputs)
    x = Dropout(0.3)(x)
    # LSTM Layer 2
    x = LSTM(128)(inputs)
    x = Dropout(0.3)(x)
    # LSTM Layer 3
    x = LSTM(64)(inputs)
    x = Dropout(0.3)(x)
    # Dense Layer
    x = Dense(64, activation = 'relu')(x)
    x = Dense(1, activation = 'sigmoid')(x)
    model = Model(inputs, x)
    # Model compiler
```

```

        model.compile(loss='binary_crossentropy',
optimizer='adam', metrics=['acc'])
        # Model Checkpoint
        mc = ModelCheckpoint('best_model_lstm.hdf5',
monitor='val_loss', verbose=1, save_best_only=True,
mode='min')
        return model, mc
# Calling the function
model, mc = lstm_model(X_train_features)
# Model Summary:
model.summary()

```

Навчання моделі, функція втрат, оцінювання продуктивності моделі та прогнозування нових аудіоданих представлено у лістингу 3.16.

Лістинг 3.16 – Застосування моделі на основі LSTM

```

# Training the model:
history=model.fit(X_train_features, Y_train, epochs=10,
callbacks=[mc], batch_size=32,
validation_data=(X_test_features, Y_test))
# load best model weights
model.load_weights('best_model_lstm.hdf5')
# summarize history for loss
plt.plot(history.history['loss'])
plt.plot(history.history['val_loss'])
plt.title('model loss')
plt.ylabel('loss')
plt.xlabel('epoch')
plt.legend(['train', 'validation'], loc='upper left')
plt.show()

```

Втрати моделі для моделі LSTM, побудованої на основі ознак часової області, становить показники, які представлені на рисунку 3.8.

					КС КРБ 123.232.00.00 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		56

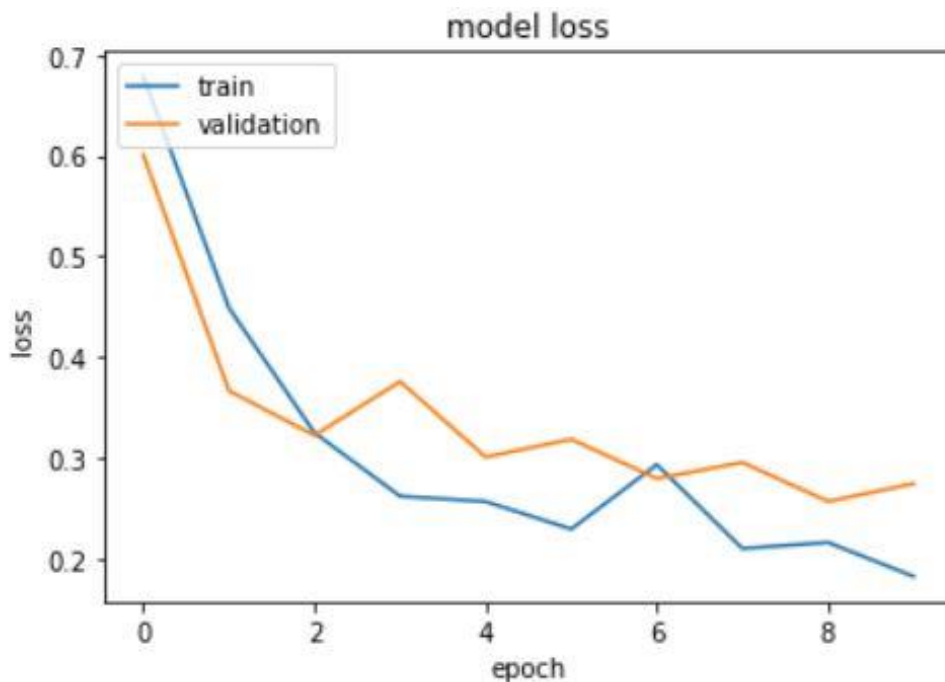


Рисунок 3.8 – Втрати моделі на основі LSTM

Оцінювання моделі та прогнозування класифікації звукових сигналів наведено у лістингу 3.17

Лістинг 3.17 – Оцінювання моделі на основі LSTM

```

# Model Evaluation:
_,acc = model.evaluate(X_test_features, Y_test)
print("Accuracy:",acc)

# For Prediction using LSTM: the input audio
index = 513
test_audio = X_test[index]

# Using IPython.display to play the audio
ipd.Audio(test_audio, rate=16000)

# classification
feature = X_test_features[index]
prob = model.predict(feature.reshape(1,-1,160))
if (prob[0][0] < 0.5):
    pred='emergency vehicle'
else:
    pred='non emergency vehicle'
print("Prediction:",pred)

```

Змн.	Арк.	№ докум.	Підпис	Дата

У результаті одержано звуковий файл із міткою, що звуковий сигнал належить звичайному автомобілю. Функція втрат зменшилася з 0,70 до 0,30, а точність підвищилася з 77% до 91% завдяки використанню LSTM над алгоритмом CNN. У даному випадку, використано ознаки часової області в обох моделях, і в подальшому їх можна використовувати для формування характеристик спектрограми.

3.3 Модель класифікації з врахуванням ознак спектрограми

У випадку з ознаками часової області вихідні аудіодані були безпосередньо опрацьовані. Однак для роботи зі спектрограмою необхідно:

- перетворити аудіодані в спектрограму;
- передайте спектрограму як вхідний масив моделі.

Спектрограма приймає неопрацьовану звукову хвилю, а потім розбиває її на послідовності або фрагменти і потім застосовує швидке перетворення Фур'є (FFT) до кожного вікна для обчислення частот.

Параметрами для обчислення спектрограми:

- параметр «Nperseg» — це розмір вікна, тобто кількість семплів у кожному аудіо;
- параметр Noverlap – це кількість реалізацій сигналу, що перекриваються між кожним вікном і розглядаються яким чином перекриваються вибірки між послідовними звуковими фрагментами.

У лістингу 3.18 показано програмну реалізацію функції формування спектрограми.

Лістинг 3.18 – Функція реалізації спектрограми

```
def spec_log(audio, sample_rate, eps = 1e-10):  
    freq, times, spec = scipy.signal.spectrogram(audio, fs=  
sample_rate, nperseg=320, noverlap=160)  
    return freq, times, np.log(spec.T.astype(np.float32) +  
eps)
```

					КС КРБ 123.232.00.00 ПЗ	Арк.
						58
Змн.	Арк.	№ докум.	Підпис	Дата		

Внаслідок реалізації функція `scipy.signal.spectrogram` одержано три масиви:

- масив «freq» — це масив частот;
- масив «times» – це час сегмента масиву;
- масив «spec» – спектрограма.

Тут важливо обчислити логарифм спектрограми, оскільки значення амплітуди можуть бути дуже малими значеннями (10^y - 10^{-10} степені) і при цьому складно аналізувати такі малі значення.

Для візуалізації спектрограм потрібні два аргументи: спектрограма та мітки для розрізнення. У лістингу 3.19 представлено функцію для візуалізації спектрограм, а на рисунку 3.9 безпосередньо їх вигляд.

Лістинг 3.19 – Функція візуалізації спектрограми та її виклику

```
def spec_plot(spectrogram, label):
    fig = plt.figure(figsize=(14,8))
    ax = fig.add_subplot(211)
    ax.imshow(spectrogram.T, aspect='auto',
    extent=[times.min(), times.max(), freqs.min(), freqs.max()])
    ax.set_title('Spectrogram of '+label)
    ax.set_xlabel('Seconds')
    ax.set_ylabel('Freqs in Hz')
    plt.show()

# Computing and Visualizing the Spectrogram for Emergency:
freqs, times, spectrogram = spec_log(emergency[162],
sample_rate)
spec_plot(spectrogram, "Emergency")

# Computing and Visualizing the Spectrogram for Non-
Emergency:
freqs, times, spectrogram = spec_log(non_emergency[162],
sample_rate)
spec_plot(spectrogram, "Non Emergency")
```

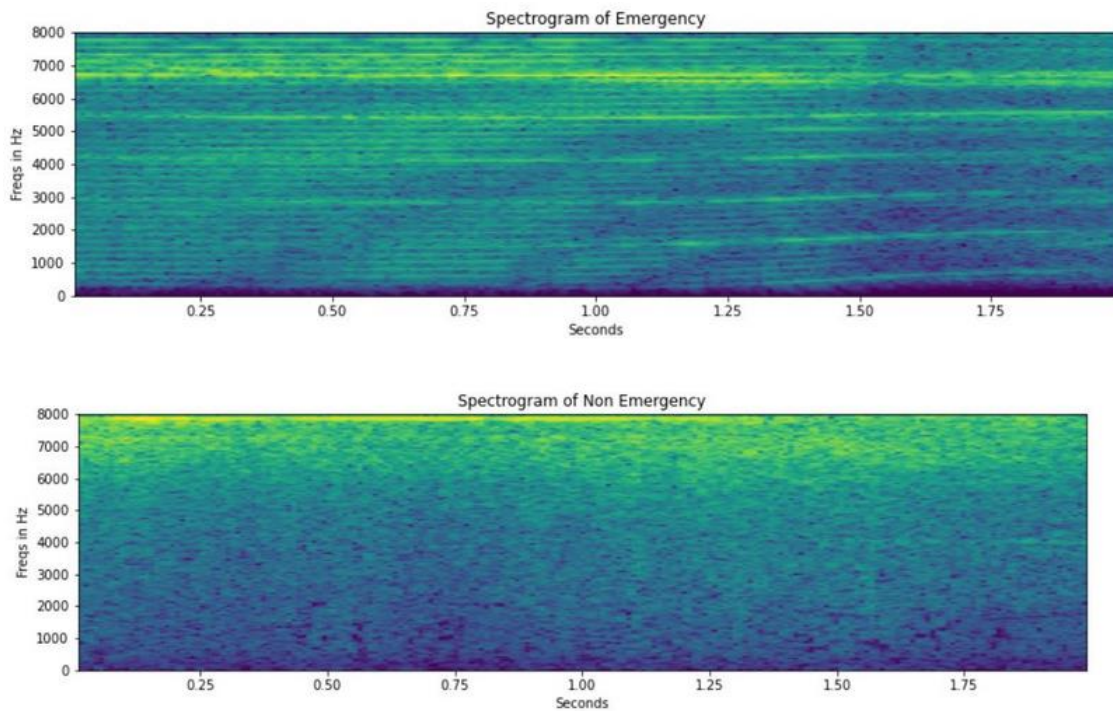


Рисунок 3.9 – Вигляд спектрограм для звукових сигналів авто екстрених служб та звичайних авто

На спектрограмі звичайного автомобіля горизонтальні лінії не зображені. Це могло бути через відсутність звуку сирени.

Наступний крок полягає у добуванні ознак спектрограми з того самого набору аудіо, який використовувався вище для ознак звукового сигналу у часовій області. Лістинг 3.20 демонструє виявлення ознак спектрограми.

Лістинг 3.9 – Добування ознак спектрограми

```
def extract_spec_features(X_tr):
    # defining empty list to store the features:
    features = []
    # We only need the 3rd array of Spectrogram so assigning
    the first two arrays as _
    for i in X_tr:
        _,_, spectrogram = spec_log(i, sample_rate)
        mean = np.mean(spectrogram, axis=0)
        std = np.std(spectrogram, axis=0)
        spectrogram = (spectrogram - mean)/std
        features.append(spectrogram)
    # returning the features as array
```

```

return np.array(features)

# Calling extract function to get training and testing sets:
X_train_features = extract_spec_features(X_train)
X_test_features = extract_spec_features(X_test)

```

Добуті ознаки згодом будуть використані для передачі алгоритму при побудові моделі. Далі потрібно навчити модель на основі CNN на тих же характеристиках спектрограми, використовуючи визначені вище функції для моделі CNN (лістинг 3.21). На рисунку 3.10 показано структуру нейронної мережі, яка використовується для навчання.

Model: "model_11"

Layer (type)	Output Shape	Param #
input_15 (InputLayer)	(None, 199, 161)	0
conv1d_20 (Conv1D)	(None, 199, 16)	28352
dropout_29 (Dropout)	(None, 199, 16)	0
max_pooling1d_20 (MaxPooling)	(None, 99, 16)	0
global_max_pooling1d_10 (GlobalMaxPooling1D)	(None, 16)	0
dense_21 (Dense)	(None, 16)	272
dense_22 (Dense)	(None, 1)	17
Total params: 28,641		
Trainable params: 28,641		
Non-trainable params: 0		

Рисунок 3.10 – Структура моделі нейронної мережі для аналізу спектрограми

Лістинг 3.21 – Навчання моделі CNN за ознаками спектрограми

```

# Model Training:
history=model_3.fit(X_train_features, Y_train, epochs=10,
callbacks=[modelcheckpoint_3], batch_size=32,

```

```

validation_data=(X_test_features,
Y_test))
# summarize history for loss
plt.plot(history.history['loss'])
plt.plot(history.history['val_loss'])
plt.title('model loss')
plt.ylabel('loss')
plt.xlabel('epoch')
plt.legend(['train', 'validation'], loc='upper left')
plt.show()
# model's performance on the test set:
_,acc = model_3.evaluate(X_test_features, Y_test)
print("Accuracy:",acc)

```

У результаті навчання нейронної мережі одержано точність 97%. Однак можна ще більше покращити точність із застосуванням різних алгоритмів, використовуючи ознаки частотної області та змінюючи довжину вхідної послідовності. На рисунку 3.11 показано структуру реалізованої нейронної мережі з використанням LSTM на ознаках спектрограми.

Model: "model_12"

Layer (type)	Output Shape	Param #
input_16 (InputLayer)	(None, 199, 161)	0
lstm_12 (LSTM)	(None, 64)	57856
dropout_32 (Dropout)	(None, 64)	0
dense_23 (Dense)	(None, 64)	4160
dense_24 (Dense)	(None, 1)	65
Total params: 62,081		
Trainable params: 62,081		
Non-trainable params: 0		

Рисунок 3.11 – Архітектура моделі нейронної мережі з використанням LSTM та ознак спектрограми

Програмний код навчання нейронної мережі на основі LSTM на основі ознак спектрограми показано у лістингу 3.22.

Лістинг 3.22 – Навчання мережі на основі LSTM та ознак спектрограми

```
# Train the model
history=model_2.fit(X_train_features, Y_train, epochs=10,
callbacks=[mc_2], batch_size=32,
validation_data=(X_test_features,
Y_test))
model_2.load_weights('best_model_lstm.hdf5')
# summarize history for loss
plt.plot(history.history['loss'])
plt.plot(history.history['val_loss'])
plt.title('model loss')
plt.ylabel('loss')
plt.xlabel('epoch')
plt.legend(['train', 'validation'], loc='upper left')
plt.show()
```

У результаті одержано графік функції втрат, який показано на рисунку 3.12, точність якої досягнуто за допомогою моделі LSTM становить 90,5%.

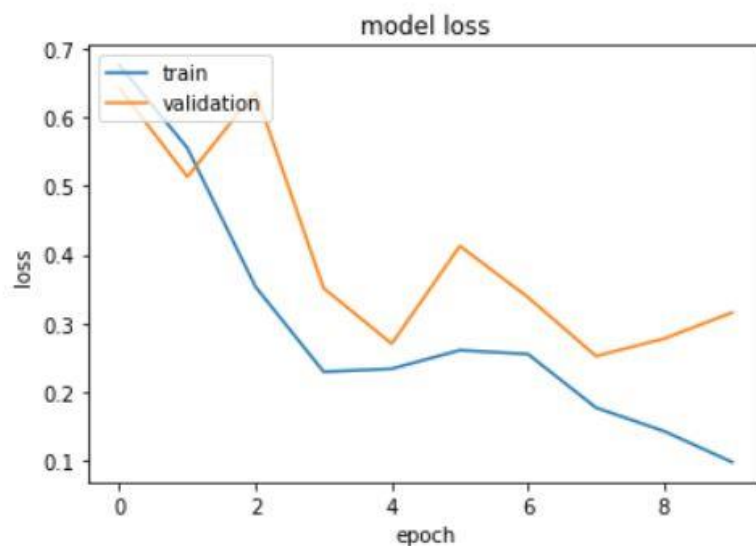


Рисунок 3.12 – Функція втрат при оцінювання моделі на основі LSTM

У даному розділі програмно реалізовано модель розпізнавання та класифікації звукових сигналів для виявлення авто екстрених служб, що в подальшому імплементовано на пристрій Arduino Nano 33 BLE Sense.

					<i>КС КРБ 123.232.00.00 ПЗ</i>	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		64

РОЗДІЛ 4 БЕЗПЕКА ЖИТТЄДІЯЛЬНОСТІ, ОСНОВИ ОХОРОНИ ПРАЦІ

4.1 Вимоги до режимів праці і відпочинку при роботі з ВДТ

При організації праці, пов'язаної з використанням ВДТ ЕОМ і ПЕОМ, для збереження здоров'я працюючих, запобігання професійним захворюванням і підтримки працездатності передбачаються внутрішньозмінні регламентовані перерви для відпочинку.

Внутрішньозмінні режими праці і відпочинку містять додаткові нетривалі перерви в періоди, що передують появі об'єктивних і суб'єктивних ознак стомлення і зниження працездатності.

При виконанні робіт, що належать до різних видів трудової діяльності, за основну роботу з ВДТ вважається така, що займає не менше 50% робочого часу. Впродовж робочої зміни мають передбачатися:

- перерви для відпочинку і вживання їжі (обідні перерви);
- перерви для відпочинку і особистих потреб (згідно з трудовими нормами);
- додаткові перерви, що вводяться для окремих професій з урахуванням особливостей трудової діяльності.

За характером трудової діяльності розрізняють три професійні групи, згідно з діючим класифікатором професій (ДК-003-95 і Зміна N1 до ДК-003-95):

- розробники програм (інженери-програмісти) виконують роботу переважно з відеотерміналом та документацією при необхідності інтенсивного обміну інформацією з ЕОМ і високою частотою прийняття рішень. Робота характеризується інтенсивною розумовою творчою працею з підвищеним напруженням зору, концентрацією уваги на фоні нервово-емоційного напруження, вимушеною робочою позою, загальною гіподинамією,

					КС КРБ 123.232.00.00 ПЗ			
<i>Змн.</i>	<i>Арк.</i>	<i>№ докум.</i>	<i>Підпис</i>	<i>Дата</i>				
<i>Розроб.</i>		Сахарчук В.В.			<i>Безпека життєдіяльності, основи охорони праці</i>	<i>Літ.</i>	<i>Арк.</i>	<i>Аркуші</i>
<i>Перевірів.</i>		Яцишин В.В.					65	
<i>Консульт.</i>		Лазарюк В.В.				<i>ТНТУ, каф. КС, гр. СІс-43</i>		
<i>Н. Контр.</i>		Луцик Н.С.						
<i>Затверд.</i>		Осухівська Г.М.						

періодичним навантаженням на кисті верхніх кінцівок. Робота виконується в режимі діалогу з ЕОМ у вільному темпі з періодичним пошуком помилок в умовах дефіциту часу;

– оператори електронно-обчислювальних машин виконують роботу, пов'язану з обліком інформації, одержаної з ВДТ за попереднім запитом, або тієї, що надходить з нього, супроводжується перервами різної тривалості, пов'язана з виконанням іншої роботи і характеризується напруженням зору, невеликими фізичними зусиллями, нервовим напруженням середнього ступеня та виконується у вільному темпі;

– оператор комп'ютерного набору виконує одноманітні за характером роботи з документацією та клавіатурою і нечастими нетривалими переключеннями погляду на екран дисплея, з введенням даних з високою швидкістю. Робота характеризується як фізична праця з підвищеним навантаженням на кисті верхніх кінцівок на фоні загальної гіподинамії з напруженням зору (фіксація зору переважно на документи), нервово-емоційним напруженням.

Правилами встановлюються такі внутрішньозмінні режими праці та відпочинку при роботі з ЕОМ при 8-годинній денній робочій зміні в залежності від характеру праці:

– для розробників програм із застосуванням ЕОМ слід призначати регламентовану перерву для відпочинку тривалістю 15 хвилин через кожну годину роботи за ВДТ;

– для операторів із застосуванням ЕОМ слід призначати регламентовані перерви для відпочинку тривалістю 15 хвилин через кожні дві години;

– для операторів комп'ютерного набору слід призначати регламентовані перерви для відпочинку тривалістю 10 хвилин після кожної години роботи за ВДТ.

У всіх випадках, коли виробничі обставини не дозволяють застосувати регламентовані перерви, тривалість безперервної роботи з ВДТ не повинна перевищувати 4 години.

					КС КРБ 123.232.00.00 ПЗ	Арк.
						66
Змн.	Арк.	№ докум.	Підпис	Дата		

При 12-годинній робочій зміні регламентовані перерви повинні встановлюватися в перші 8 годин роботи аналогічно перервам при 8-годинній робочій зміні, а протягом останніх 4-х годин роботи, незалежно від характеру трудової діяльності, через кожен годину тривалістю 15 хвилин.

Для зниження нервово-емоційного напруження, втомлення зорового аналізатора, поліпшення мозкового кровообігу, подолання несприятливих наслідків гіподинамії, запобігання втомі доцільно деякі перерви використовувати для виконання комплексу вправ, які наведені у Державних санітарних правилах і нормах роботи з візуальними дисплейними терміналами електронно-обчислювальних машин ДСанПІН 3.3.2.007-98.

4.2 Вплив електромагнітних полів (ЕМП) на людину та заходи щодо зменшення їх впливу на обслуговуючий персонал

Сучасний рівень розвитку техніки означає високий рівень використання різноманітної радіоелектронної апаратури. В основному всю радіоапаратуру можна розділити на дві великі групи: засоби зв'язку і засоби обробки та зберігання даних в найрізноманітнішому вигляді.

Основну небезпеку для людини становить високочастотне електромагнітне випромінювання. Вплив випромінювання низької частоти вважається безпечним, хоча існують версії, що вплив випромінювання низької частоти проявляється через значно довший період, співрозмірний із тривалістю життя людини.

Основними джерелами електромагнітного випромінювання в електронно-обчислювальній техніці є генератори тактової частоти, сигнали яких використовуються для синхронізації роботи електронних схем обчислювальної техніки. Випромінювання цих пристроїв є малопотужним, хоча і високочастотним. Низький рівень випромінювання диктується тим, щоби мінімізувати взаємний вплив сигналів одних кіл на інші, а, отже, і мінімізувати спотворення.

					КС КРБ 123.232.00.00 ПЗ	Арк.
						67
Змн.	Арк.	№ докум.	Підпис	Дата		

Більш потужний електромагнітний фон випромінюють блоки живлення обчислювальної техніки, оскільки більшість з них – імпульсні. Але, як правило, блоки живлення не розміщені в безпосередній близькості до користувача і тому не становлять особливої небезпеки для людини.

Системний блок комп'ютера не несе особливої небезпеки для життя та здоров'я користувача. Найбільшу небезпеку несе підвищене значення напруги, що подається з електромережі на блок живлення системного блоку. Несприятливий вплив на користувача може здійснювати шум, що створюється при роботі вентиляторів та накопичувачів системного блоку.

Електромагнітні випромінювання, які виникають при роботі електронних компонентів блоку мають незначні рівні, тому можуть створювати хіба що радіочастотні перешкоди.

Більшість електронних компонентів комп'ютера живляться напругою "+5 В", двигуни накопичувачів "+12 В", деякі пристрої "- 5 В" та "- 12 В". Такі значення напруги не несуть небезпеки для людини. В той же час на вхід блоку живлення подається напруга електромережі (220 В), яка є небезпечною для людини, з точки зору її можливого ураження. Тому до блоку живлення висувається низка вимог електробезпеки. Зокрема, електропроводи та кабелі повинні мати надійну ізоляцію, а на випадок короткого замикання чи інших аварійних режимів в електричній схемі блоку живлення повинні бути передбачені елементи захисту. Сам же він знаходиться в корпусі, який перекриває доступ до струмоведучих елементів блоку.

Дисплей є одним з основних блоків комп'ютера, що призначений для візуалізації інформації. Від його характеристик в значній мірі залежить працездатність та стан здоров'я користувача комп'ютера.

Концепція нормування електромагнітних полів і випромінювань передбачає:

- вироблення єдиної системи нормативних значень гранично допустимих рівнів електромагнітних полів і випромінювань;
- захист природних ресурсів від втрат, обумовлених дією цих полів на різні компоненти природного середовища;

					КС КРБ 123.232.00.00 ПЗ	Арк.
						68
Змн.	Арк.	№ докум.	Підпис	Дата		

– запобігання значним функціональним порушенням екосистем в результаті прямої або непрямой дії полів на ті або інші компоненти цих систем.

Монітор і ПЕОМ повинні забезпечувати потужність експозиційної дози рентгенівського випромінювання в будь-якій точці на відстані 0,05м від екрана і корпусу монітора при будь-яких положеннях регульовальних пристроїв не вище $7,74 \times 10$ А/кг, що відповідає еквівалентній дозі, рівної 0,1 мбер/година (100 мкр/годину).

Дисплеї, сконструйовані на основі ЕПТ, є джерелом кількох видів електромагнітного випромінювання та полів, а саме:

- іонізуючого (рентгенівського) випромінювання;
- оптичного випромінювання;
- електромагнітних випромінювань та полів.

Джерелом рентгенівського випромінювання дисплею є ЕПТ, зокрема люмінофорне покриття екрана. Електрони, які летять з великою швидкістю, різко сповільнюються шаром люмінофора. Це й спричинює рентгенівське випромінювання, енергія якого обмежується потенціалом розгону електронів. При цьому можлива поява лише, так званого "м'якого" рентгенівського випромінювання з довжиною хвилі 2 - 5 нм, яке ефективно поглинається скляним екраном.

Оптичне випромінювання виникає в результаті взаємодії електронів з шаром люмінофору. Дослідження показали, що в процесі роботи дисплея, окрім видимого випромінювання, мають місце і інші види оптичного-випромінювання (ультрафіолетове та інфрачервоне).

Електромагнітні випромінювання та поля різних діапазонів частот (високих, низьких та вкрай низьких) виникають в системах горизонтальної та вертикальної розгортки та в результаті дії електронного променя.

Основним джерелом статичних електричних полів є висока напруга (6 - 15 кВ), яка подається на блок анодів та внутрішню поверхню екрана ЕПТ.

Реальна інтенсивність, напруженість, рівень та інші параметри кожного виду електромагнітного випромінювання залежить від технічної конструкції конкретного дисплея, режимів його роботи, екранування та інших факторів.

					КС КРБ 123.232.00.00 ПЗ	Арк.
						69
Змн.	Арк.	№ докум.	Підпис	Дата		

Центральним елементом лазерного принтера є експонуючий барабан. Його поверхня покрита, як правило, органічним світлочутливим непровідним матеріалом. Спочатку поверхня експонуючого барабана заряджається від'ємним зарядом. Далі в дію вступає лазерний промінь, що відтворює інформацію, отриману принтером з комп'ютера. В місцях, на які потрапив промінь лазера, відбувається зміна знаку електричного заряду. За допомогою валиків подачі і перенесення тонера, останній потрапляє на експонуючий барабан. Тонер прилипає лише в місцях, що були опромінені лазером. З касети до барабану подається папір. Транспортуючий валик притискає його до барабана, при цьому тонер переходить на папір.

Вузол фіксування складається, частіше за все, з двох барабанів, один із яких нагрівається лампою до температури близько 200 °С. Під дією температури тонер розплавляється і назавжди зв'язується з папером. Далі віддрукований аркуш подається у приймальний відділ принтера. Надлишковий тонер, що не перейшов на папір, забирається з барабана у вузлі очистки. В кінці циклу світлочутливе покриття барабана розряджається за допомогою лампи.

Таким чином, розглянувши роботу лазерного принтера, можна виділити такі його шкідливі та небезпечні фактори за порядком їх проявлення в розглянутому вище процесі друкування:

- електростатичні заряди;
- лазерний промінь;
- висока температура.

Для захисту людини від шкідливого впливу електромагнітних полів приймаються нормативи, стандарти та норми, які пов'язані із захистом людини від небезпечного впливу. Вони завжди представляють собою компроміс між перевагами використання нових технологій і техніки та можливим ризиком, спричиненим їхнім використанням.

ГОСТ 12.1.006-84 “Електромагнітні поля радіочастот” охоплює діапазон частот 60 кГц-300МГц. Він встановлює, що оцінка ЕМП в діапазоні 60 кГц-300МГц проводиться окремо з електричних і магнітних складових поля. Допустимі рівні протягом робочого дня по електричній складовій не повинні

					КС КРБ 123.232.00.00 ПЗ	Арк.
						70
Змн.	Арк.	№ докум.	Підпис	Дата		

перевищувати 50 В/м знижуючись ступенями 5 В/м на міру підвищення частоти. По магнітній складовій встановлені рівні тільки для окремих ділянок діапазону: 5 А/м для частот 60 кГц-1.5 МГц та 0,3 А для частот 30-50 МГц. Допускається перевищення цих стандартів, але не більше ніж двократно, при скороченні робочого дня не менш як на 50% [14].

Для частот 300 МГц-30 ГГц гранично допустимі значення щільності визначаються як результат ділення нормованої величини енергетичного навантаження за робочий день на час впливу. Енергетичне навантаження протягом робочого дня не повинно перевищувати 200 мкВт год./ см².

Електромагнітні поля сильно впливають на людський організм. Вони негативно впливають майже на усі системи організму. Тому треба створювати певні методи захисту від їх дії. Найпоширенішими методами є [14]:

- зменшення щільності потоку енергії, якщо дозволяє даний технологічний процес або обладнання;
- захист часом (тобто обмеження часу знаходження у зоні джерела ЕМП);
- захист відстанню;
- екранування робочого місця чи джерела;
- раціональне планування робочого місця;
- застосування засобів попереджувальної сигналізації;
- застосування засобів особистого захисту.

Для зменшення впливу електромагнітних полів на персонал, який знаходиться у зоні дії деяких радіоелектронних засобів необхідним є ряд захисних заходів: організаційні, інженерно-технічні та лікувально-профілактичні.

На етапі проектування системи взаємне розміщення об'єктів має бути забезпечено таким чином, щоб інтенсивність опромінення була мінімальною. Окрім цього, потрібно також зменшити час перебування персоналу у зоні опромінення. Потужність джерел випромінювання має бути найменшою з можливих.

					КС КРБ 123.232.00.00 ПЗ	Арк.
						71
Змн.	Арк.	№ докум.	Підпис	Дата		

ВИСНОВКИ

При виконанні кваліфікаційної роботи досягнуто поставленої мети щодо проектування комп'ютерної системи автоматичного розпізнавання і класифікації звукових сигналів на основі методів і засобів спектрального аналізу дискретних звукових сигналів та параметрів його амплітуди і частоти.

Комп'ютерна система на апаратному рівні використовує плату Arduino Nano 33 BLE Sense, що містить в собі ряд інтегрованих сенсорів, серед яких мікрофон підвищеної чутливості з шумопоглинанням. Це дає змогу одержувати звуковий сигнал в реальному часі і за допомогою реалізованої моделі інтелектуального розпізнавання і класифікації аудіо сигналів розпізнавати звуки сирен автомобілів екстрених служб.

Інтелектуальну систему розпізнавання і класифікації звукових сигналів реалізовано за допомогою мови програмування Python та відкритих бібліотек. Після цього дану модель завантажено в оперативну пам'ять Arduino Nano. Точність розпізнавання звукових сигналів становить вище 90%, а у деяких випадках досягає 97%.

Стійкість реалізованого алгоритму становить більше, ніж 0.7. При реалізації інтелектуальної складової використано підхід глибоких нейронних мереж, зокрема згорткові нейронні мережі та LSTM.

Поряд з цим у роботі досліджено та розв'язано задачі пов'язані з аналізом природи і параметрами звукових сигналів, методами і засобами перетворення аналогових сигналів у цифрові формат, обґрунтованістю використаного апаратного і програмного забезпечення.

					КС КРБ 123.232.00.00 ПЗ	Арк.
						72
Змн.	Арк.	№ докум.	Підпис	Дата		

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Фоусет Т., Провост Ф. Data Science для бизнеса: Як збирати, аналізувати і використовувати дані. Наш формат. Київ. 2019. 400 с.
2. К. О Нил, Шатт Р. Data Science. Инсайдерская информация для новичков. Включая язык R. Издательский дом "Питер". 2018. 368 с.
3. Френкс Л. Теория сигналов. М.: Сов. радио, 1974. – 344 с.
4. Оппенгейм А., Шафер Р. Цифровая обработка сигналов. М.: Техносфера, 2006. 856 с.
5. Айфичер Э.С., Джервис Б.У. Цифровая обработка сигналов: практический подход. М.: Издательский дом «Вильямс», 2004. 992 с.
6. Марпл-мл. С.Л. Цифровой спектральный анализ и его приложения: Пер. с англ. М.: Мир, 1990. 548 с.
7. Робинсон Э.А. История развития теории спектрального оценивания. / ТИИЭР. 1982. т.70, No 9. С. 6 – 32.
8. Лайонс Р. Цифровая обработка сигналов. М.: ООО «Бином-пресс», 2006. 656 с.
9. Сергиенко А.Б. Цифровая обработка сигналов: Учебник для вузов. 2-е изд. СПб.: Питер, 2007. 751 с.
10. Linden G., Smith B., York J. Amazon.com recommendations: Item-to-item collaborative filtering. IEEE Internet Computing. vol. 7. No. 1. 2003. pp. 76–80.
11. Python-recsys on Github. URL: <https://github.com/ocelma/python-recsys> (дата звернення 22.04.2022 р).
12. Preprocessing data. URL: <https://scikit-learn.org/stable/modules/preprocessing.html#preprocessing> (дата звернення 02.05.2022 р.).
13. API reference. URL: <https://pandas.pydata.org/docs/reference/index.html> (дата звернення 10.05.2022 р.).
14. NumPy Reference. URL: <https://numpy.org/doc/stable/reference/index.html> (дата звернення 12.05.2022 р.)
15. Барсегян А. Анализ данных и процессов. 3 изд. БХВ-Петербург. 2009. 512 с.

					КС КРБ 123.232.00.00 ПЗ	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		73

16. Breese J., Heckerman D., Kadie C. Empirical Analysis of Predictive Algorithms for Collaborative Filtering. Proc. 14th Conf. Uncertainty in Artificial Intelligence. 1998. pp. 223-234.

17. Kurucz M., Benczur A., Csalogany K. Methods for large scale SVD with missing values. Proceedings of KDD Cup and Workshop. 2007. pp. 122-129.

18. ДСанПіН 3.3-2.007-98 Державні санітарні правила і норми роботи з візуальними дисплейними терміналами електронно-обчислювальних машин. - Київ, 1999. - 18с.

19. НПАОП 0.00-1.28-10 «Правила охорони праці під час експлуатації електронно-обчислювальних машин». – Київ, 2010. – 8 с.

					<i>КС КРБ 123.232.00.00 ПЗ</i>	Арк.
						74
Змн.	Арк.	№ докум.	Підпис	Дата		

Додаток А.
Технічне завдання

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Тернопільський національний технічний університет імені Івана Пулюя

Факультет комп'ютерно-інформаційних систем і програмної інженерії

Кафедра комп'ютерних систем та мереж

“Затверджую”

Завідувач кафедри КС

_____ Осухівська Г.М.

“ ____ ” _____ 2022 р

КОМП'ЮТЕРНА СИСТЕМА АВТОМАТИЧНОГО РОЗПІЗНАВАННЯ ТА
КЛАСИФІКАЦІЇ ЗВУКОВИХ СИГНАЛІВ

ТЕХНІЧНЕ ЗАВДАННЯ

на 13 листках

Вид робіт:

Кваліфікаційна робота

На здобуття освітнього ступеня «Бакалавр»

Спеціальність 123 «Комп'ютерна інженерія»

«УЗГОДЖЕНО»

«ВИКОНАВЕЦЬ»

Керівник кваліфікаційної роботи

Студент групи СІс-43

_____ к.т.н., доц. Яцишин В.В.

_____ Сахарчук В.В.

« ____ » _____ 2022 р.

« ____ » _____ 2022 р.

Тернопіль 2022

1 Загальні відомості

1.1 Повна назва та її умовне позначення

Повна назва теми кваліфікаційної роботи: «Комп'ютерна система автоматичного розпізнавання та класифікації звукових сигналів».

Умовне позначення кваліфікаційної роботи: КС КРБ 123.235.00.00

1.2 Виконавець

Студент групи СІс-43, факультету комп'ютерно-інформаційних систем і програмної інженерії, кафедри комп'ютерних систем та мереж, Тернопільського національного технічного університету імені Івана Пулюя, Сахарчук Вадим Вікторович.

1.3 Підстава для виконання роботи

Підставою для виконання кваліфікаційної роботи є наказ по університету (№ 4.7-97 від 10.02.2022 р.)

1.4 Планові терміни початку та завершення роботи

Плановий термін початку виконання кваліфікаційної роботи – 10.02.2022 р.

Плановий термін завершення виконання кваліфікаційної роботи – 23.06.2022 р.

1.5 Порядок оформлення та пред'явлення результатів роботи

Порядок оформлення пояснювальної записки та графічного матеріалу здійснюється у відповідності до чинних норм та правил ІСО, ГОСТ, ЕСКД, ЕСПД та ДСТУ.

Пред'явлення проміжних результатів роботи з виконання кваліфікаційної роботи здійснюється у відповідності до графіку, затвердженого керівником роботи.

Попередній захист кваліфікаційної роботи відбувається при готовності роботи на 90% , наявності пояснювальної записки та графічного матеріалу.

Пред'явлення результатів кваліфікаційної роботи відбувається шляхом захисту на відповідному засіданні ЕК, ілюстрацією основних досягнень за допомогою графічного матеріалу.

2 Призначення і цілі створення системи

2.1 Призначення системи

Комп'ютерна система розпізнавання та класифікації звукових сигналів призначена для автоматизації процесу визначення і встановлення приналежності згенерованих джерелом звуку сигналів на основі яких приймаються певні рішення.

Проектована система повинна забезпечувати можливість збору звукових сигналів і перетворення їх у цифровий формат. Окрім цього, одержані сигнали з відповідними характеристиками на основі моделей та алгоритмів машинного навчання повинні бути автоматично прокласифіковані за відповідними класами з відповідними мітками.

Комп'ютерні системи розпізнавання і класифікації звукових сигналів мають широкий спектр застосування, починаючи від автономних незалежних систем, закінчуючи підсистемами більш комплексних ІТ-рішень.

Автономно такі системи можуть бути використані при формуванні альбомів музики за жанрами, виконавцями чи іншими категоріями. В якості підсистем більш складних рішень, комп'ютерна система розпізнавання і класифікації звукових сигналів може бути використана при організації безпеки та голосового управління «розумного будинку».

Окрім цього, зважаючи на кількість автомобілів і їх постійне зростання при незмінній інфраструктурі, актуальним є створення систем аналізу звукових сигналів автомобілів. Це дозволить екстреним службам більш швидко та ефективно добиратись до місць виникнення небезпечних ситуацій, шляхом формування зелених коридорів.

Комп'ютерна система повинна дозволяти накопичувати дані про звукові сигнали автомобілів та автоматично встановлювати мітки приналежності звукового сигналу. На апаратному рівні система повинна бути швидкореагуючою на події, переносимою і здатною співіснувати з іншими сумісними системами.

2.2 Мета створення системи

Мета проектування та реалізації комп'ютерної системи розпізнавання та класифікації звукових сигналів полягає у створенні ефективного рішення як на апаратному, так і на програмному рівні, яка б дозволяла проводити бінарну класифікацію і прогнозування звукових сигналів для екстрених служб і формування пріоритету їх проїзду на перехрестях із світлофорами.

Передбачається, що досягнення мети роботи можливе при розв'язанні наступних задач:

- аналіз природи і параметрів звукових сигналів;
- аналіз методів і засобів перетворення аналогових звукових сигналів у цифровий формат;
- обґрунтування апаратного забезпечення при проектуванні комп'ютерної системи;
- аналіз програмних засобів для роботи із звуковими сигналами та моделей машинного навчання для їх автоматичного розпізнавання та класифікації;
- реалізація програмних моделей класифікації аудіосигналів з визначеною точністю і достовірністю;
- аналіз результатів проектування і реалізації комп'ютерної системи.

2.3 Характеристика об'єкту

2.3.1 Основні задачі та функції об'єкту

До основних задач і функцій системи розпізнавання і класифікації звукових сигналів належить здатність до накопичення і зберігання аудіо даних, автоматичної класифікації сигналів за визначеними класами і нанесення відповідних міток на прогнозовані сигнали.

Реалізацію комп'ютерної системи необхідно виконати із застосуванням мікроконтролерів з підтримкою та опрацювання алгоритмів машинного навчання. Продуктивність системи повинна бути високою для забезпечення зелених коридорів спецмашинам у заторах та на перехрестях зі світлофорами.

Система повинна володіти характеристиками щодо мінімальності габаритних розмірів і високої чутливості мікрофона з можливістю подавлення зовнішніх шумів.

Найбільш критичними задачами, які повинна розв'язувати комп'ютерна система є автоматична класифікація звукових сигналів у визначених часових рамках. Окрім цього, точність розпізнавання сирен автомобілів екстрених служб повинна бути на рівні не менше, ніж 90 %.

Додатковими задачами, які потрібно вирішити поряд з основною є наступні:

- апаратна та програмна узгодженість компонентів комп'ютерної системи;
- наявність можливості функціонування моделі класифікації звукових сигналів на мобільних і портативних пристроях, в основі якого лежить мікроконтролер;
- застосування звукових даних, які інтерпретують сирени спец автомобілів, з відкритих джерел;
- використання швидких та робастних алгоритмів при бінарній класифікації звуків сирен;
- забезпечення гнучкості налаштування параметрів при навчанні моделі розпізнавання звукових сигналів.

Найбільш раціональним є використання апаратного забезпечення з вбудованими сенсорами звуку.

Технічні характеристики апаратних засобів для проведення моделювання і тестування інтелектуальної складової комп'ютерної системи повинні підтримувати механізми розпаралелювання і мати достатній об'єм оперативної пам'яті.

3 Вимоги до системи

3.1 Вимоги до системи в цілому

До комп'ютерної системи розпізнавання і класифікації звукових сигналів в цілому висувають наступні вимоги:

- вимоги до чутливості і дальності розпізнавання сирен автомобілів екстрених служб;
- вимоги до швидкодії, оперативної пам'яті, живлення і габаритних розмірів мікроконтролера керування процесом розпізнавання звукових сигналів ;
- вимоги до навчальної і тестової вибірки звукових сигналів при побудові моделі автоматичної класифікації сигналів;
- вимоги продуктивності та зручності налаштування параметрів моделей розпізнавання звукових сигналів.

Обґрунтування вибору мікрофону зумовлено факторами його розміру, типом з'єднання з мікроконтролером та підвищеною чутливістю з шумопоглинальним фільтром.

Вимогами до мікроконтролера є достатність системних параметрів, зокрема тактової частоти процесора, його розрядності та оперативної пам'яті для швидкої видачі результатів щодо одержаного звукового сигналу в режимі реального часу.

Програмна складова комп'ютерної системи повинна забезпечувати сумісність з особливостями апаратного забезпечення та зручність використання та налаштування моделей розпізнавання і класифікації звукових сигналів.

Основні вимоги до комп'ютерної системи:

- можливість одержання аналогового звукового сигналу і перетворення його у цифровий;
- наявність аудіо файлів із записами звукових сигналів сирен автомобілів екстрених служб;
- можливість навчання та донавчання моделі розпізнавання і класифікації звукових сигналів;
- здатність налаштування чутливості приймача звукових сигналів;
- здатність до інтеграції з системами керування світлофорами;
- забезпечення надійності та захищеності апаратного і програмного забезпечення формування міток звукових сигналів.

3.1.1 Вимоги до структури та функціонування системи

Структура комп'ютерної системи автоматичного розпізнавання і класифікації звукових сигналів повинна включати в себе апаратну і програмну складову.

В якості апаратного забезпечення пропонується використовувати плати мікроконтролерів з вбудованим мікрофоном та підтримкою зв'язку на основі технології Bluetooth, а також здатністю виконання програмних моделей чи підтримкою платформ машинного навчання, наприклад, TensorFlow Lite.

Програмне забезпечення комп'ютерної системи повинно реалізовувати функції розпізнавання та автоматичної класифікації аудіо сигналів на основі попередньо навчених моделей та алгоритмів.

Комп'ютерна система повинна відповідати вимогам надійності до такого класу систем, а також бути захищеною на різних рівнях взаємодії компонентів.

3.1.2 Вимоги до способів та засобів зв'язку між компонентами системи

Вимогами до способів і засобів зв'язку є здатність здійснювати обмін даними за допомогою енергозберігаючих протоколів Bluetooth у випадку взаємодії із

зовнішнім ПК або суміжними системами. Сама комп'ютерна система на рівні апаратного забезпечення повинна бути монолітною.

Передача даних від мікроконтролера до смартфона та сховищ даних повинна використовувати на транспортному рівні стек протоколів TCP/IP. Окрім цього, компоненти комп'ютерної системи мають бути захищені на фізичному і програмному рівні.

3.1.3 Вимоги по діагностуванню системи

Діагностику компонентів комп'ютерної системи розпізнавання і класифікації звукових сигналів можна поділити на дві частини: апаратну і програмну. Вимоги до діагностики апаратного забезпечення включають:

- здатність само ініціалізації сенсорів;
- ініціалізація мікроконтролерів;
- ідентифікація з'єднань з суміжними системами;

Діагностика програмного забезпечення проводиться згідно регламенту та у випадку виявлення порушень і помилок при його роботі. Діагностика інтелектуального модулю відбувається у випадку зниження точності прогнозів при класифікації звукових сигналів у режимі реального часу.

3.1.4 Перспективи розвитку, модернізація системи

Перспективами розвитку комп'ютерної системи є її впровадження та/або інтеграція із суміжними системами регіонального та загальнонаціонального рівня у сфері управління транспортними потоками та підвищення ефективності функціонування екстрених служб. Модернізація системи можлива при зміні вимог до її експлуатації, або у випадку побудови більш оптимальної моделі за точністю та одночасно за кількістю використовуваних ресурсів.

Перспективи розвитку і модернізації комп'ютерної системи можуть бути також пов'язані з необхідністю розширення функціональних властивостей або переходу на

іншу програмно-апаратну платформу. При цьому проєктована система повинна бути гнучкою і масштабованою.

3.1.5 Вимоги до надійності системи

Вимогами надійності до комп'ютерної системи розпізнавання та класифікації звукових сигналів виступають:

- безвідмовність функціонування протягом визначеного періоду часу;
- можливість використання однотипних та взаємозамінних сенсорів з похибками в межах допустимих інтервалів;
- безвідмовність роботи мікроконтролера Arduino Nano 33 BLE Sense;
- час напрацювання компонентів системи на відмову до 10000 год.;
- сумісність із зовнішніми однорідними системами;
- захищеність доступу до апаратного і програмного забезпечення;
- захист апаратних пристроїв від впливу агресивних зовнішніх метеофакторів.

3.1.6 Вимоги до функцій та задач, які виконує система

Основними вимогами до функцій і задач, які має розв'язувати комп'ютерна система автоматичного розпізнавання та класифікації звукових сигналів є:

- можливість одержання звукових сигналів;
- здатність перетворення аналогових характеристик мікрофона у цифровий формат;
- можливість комунікації компонентів системи за визначеними протоколами передачі даних;
- здатність накопичувати інформацію про звукові сигнали авто екстрених служб та звичайних авто;
- здатність до прогнозування міток аудіофайлів, отриманих в реальному часі;
- врахування факторів впливу на тип звукового сигналу і його характеристики;

- забезпечення точності результатів прогнозування не нижче 90% та стійкості алгоритму;

- можливість налаштування прав доступу до ресурсів комп'ютерної системи;

3.1.7 Вимоги до апаратного забезпечення

Апаратне забезпечення комп'ютерної системи автоматичного розпізнавання і класифікації звукових сигналів включає в себе Arduino Nano 33 BLE Sense з будованими сенсорами мікрофону, температури і вологості, атмосферного тиску, і давача руху.

При реалізації інтелектуального модуля для розпізнавання і класифікації звукових сигналів необхідно застосування комп'ютерної техніки з наступними технічними характеристиками:

- процесор з кількістю логічних ядер не менше 16 і тактовою частотою на рівні 2,1 ГГц;

- відеокарта об'ємом 8 ГБ;

- розмір оперативної пам'яті – не менше 32 ГБ;

- об'єм накопичувачу – 4 ТБ.

3.1.8 Вимоги до програмного забезпечення

Реалізація інтелектуального модуля повинна бути виконана з використанням мови програмування Python і відповідних відкритих бібліотек.

4 Вимоги до документації

Документація повинна відповідати вимогам ЄСКД та ДСТУ

Комплект документації повинен складатись з:

- пояснювальної записки;

- графічного матеріалу:

1. Загальна характеристика звукових сигналів

2. Схема електрична принципова Arduino Nano 33 BLE Sense.

3. Алгоритм реалізації програмного забезпечення класифікації звукових сигналів..

4. Спектрограма звукових сигналів автомобілів екстрених служб та звичайних авто.

5. Структура моделі та результати класифікації звукових сигналів

5 Стадії та етапи проектування

Таблиця 1 – Стадії та етапи виконання кваліфікаційної роботи бакалавра

№ етапу	Назва етапу виконання кваліфікаційної роботи	Термін виконання
1.	Розробка та аналіз технічного завдання	23.03-30.03.2022
2.	Аналіз принципів організації комп'ютерних систем розпізнавання і класифікації звукових сигналів	30.03-06.04.2022
3.	Обґрунтування вибору апаратного забезпечення комп'ютерної системи	07.04-15.04.2022
4.	Проектування архітектури системи розпізнавання і класифікації звукових сигналів	15.04-29.04.2022
5.	Аналіз параметрів звукових сигналів у цифровому форматі	30.04-12.05.2022
6.	Реалізація програмної моделі інтелектуального модуля	13.05-18.05.2022
7.	Розробка інструкцій з налаштування параметрів комп'ютерної системи	18.05-25.05.2022
8.	Безпека життєдіяльності, основи охорони праці	25.05-10.06.2022
9.	Оформлення кваліфікаційної роботи	10.06-12.06.2022
10.	Попередній захист кваліфікаційної роботи	13.06-18.06.2022

6 Додаткові умови виконання кваліфікаційної роботи

Під час виконання кваліфікаційної роботи у дане технічне завдання можуть вноситися зміни та доповнення.